REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ORGANIZATION.

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
26-02-2010	Final Report	12/2006-1/2009

4. TITLE AND SUBTITLE
Human Supervision of Time Critical Control Systems

5a. CONTRACT NUMBER

5b. GRANT NUMBER
AFOSR FA9550-07-1-0047

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Pardalos, Panos M.
Uryasev, Stanislav
Ding, Mingzhou
Krokhmal, Pavlo
Cannon, Jordan
Murphcy, Robert

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
University of Florida, 219 Grinter Hall, PO Box 115500, Gainesville, FL 32611-5500
University of Iowa, 105 Jessup Hall, Iowa City, IA 52242-1316

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Air Force Office of Scientific Research
875 N Randolf Street
Arlington, VA 22023

10. SPONSOR/MONITOR'S ACRONYM(S)
AFOSR

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release

13. SUPPLEMENTARY NOTES
The most significant results obtained throughout this project are represented in the archival publications. The enclosed addendum to final report presents the most recent findings that have not been, at the time of reporting, published or submitted for publication.

14. ABSTRACT
Data collection of real-time operator performance under varying levels of workload and stress conditions has been performed at AFRL/RHCP. The data includes psychophysiological indicators (5 electroencephalogram (EEG) channels, vertical electrooculogram (VEOG), horizontal electrooculogram(HEOG), and electrocardiogram (ECG)) from 12 participants who monitored a simulated mission involving several unmanned aerial vehicles (UAVs). Substantial part of research efforts was dedicated to studying and modeling of information flow in brain networks, extraction of patterns from brain signals in real time. A number of algorithms for quantitative analysis of psychophysiological data have been developed, including approaches based on p-order conic programming, support vector machines, non-parametric statistical analysis via Granger causality, etc. The conducted studies indicate that the methods based on temporal trend detection in dimensionally reduced time series, obtained via projection of Kullback-Leibler divergence, as well as independent component analysis, possess substantial robustness and can serve as predictive metrics for implementation on closed-loop architectures.

15. SUBJECT TERMS
Human supervision, time-critical system, online algorithms, psychophysiological data, cognition mining, pattern recognition, optimization, cooperative control

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			19b. TELEPHONE NUMBER *(Include erea code)*

Addendum to Final Report

AFOSR Grant FA9550-07-1-0047
Human Supervision of Time Critical Control Systems

Jordan Cannon
Pavlo Krokhmal
Robert Murphey
Panos Pardalos

Abstract. This addendum reports findings of the PIs in the course of the project "Human Supervision of Time-Critical Control Systems", which, at the moment of the report, have not been published or submitted for publication. In particular, it presents a robust technique for detecting temporal changes in multidimensional time series data represented by phychophysiological measurements. The proposed approach relics on detecting trends in an appropriate statistics obtained using independent component analysis of the data. In comparison to the previous results obtained by the PIs in the scope of this project, the new method performs much more robustly in terms of between-subject and between-trial differences. The proposed algorithm is amenable to efficient implementation compliant with on-line performance requirements.

1. Introduction

The inereasing eomplexity and sophistieation of eomputing, sensing, and communication teehnologies paves the way for proliferation of unmanned autonomous systems and platforms, which will replace and/or assist humans in hazardous or resouree consuming missions. Yet, the common consensus is that despite the ever growing need for increased autonomy of various unmanned systems and vehicles, the human supervision of sueh systems is indispensable and eritieal for mission suecess.

On the other hand, the large amounts of information produeed by these eomplex systems can place a high demand on a human operator's cognitive load, potentially overwhelming him/her and leading to degraded performance. Thus, the ability of an automated control system to estimate the current functional state of a human operator/supervisor and supply information that is eonducive to human decision-making at the given cognitive level is essential to robust and sueeessful system performanee.

The objective of the present endeavor is to advance methods for real-time detection of changes in human operator's funetional state that meet real-time or online requirements, and which can be utilized in elosed-loop autonomous control systems with human supervision. Generally, operator funetional state (OFS) can be defined as the momentary ability of an operator to meet task demands with their cognitive and physiologieal resources. In the context of this work, the OFS is assoeiated with eognitive load experienced by the operator; with this eaveat in mind, we will use both terms interchangeably.

With the ability to measure and deteet changes in OFS in real-time, a elosed-loop system between the operator and machine eould optimize OFS through the dynamie allocation of tasks. For instance, if the system detects the operator is in cognitive overload, it ean automate eertain tasks allowing them to better foeus on salient information. Conversely, if the system detects under-vigilance, it ean alloeate tasks back to the manual control of the operator. In essenee, this system operates to "dynamieally match task demands to [an] operator's momentary cognitive state", thereby aehieving optimal OFS (Wilson, Russell, 2007).

OFS is eommonly measured indirectly, like using overt performanee metrics on tasks; if performanee is deelining, a low OFS is assumed. Another indirect measure is the subjeetive estimate of mental workload, where an operator narrates his/her perceived functional state while performing tasks (Wilson, Russell, 2007). Unfortunately, indirect measures of OFS are often infeasible in operational settings; performanee metries are diffieult to construet for highly-automated eomplex systems, and subjective workload estimates are often inaccurate and intrusive (Wilson, Russell, 2007; Prinzel et al., 2000; Smith et al., 2001).

OFS can be more directly measured via psychophysiological signals sueh as eleetroeneephalogram (EEG) and eleetrooculography (EOG). Current researeh has demonstrated these signals' ability to respond to ehanging cognitive load and to measure OFS (Wilson, Fisher, 1991; Wilson, Fisher, 1995; Gevins et al., 1997; Gevins et al., 1998; Byrne, Parasuraman, 1996). Moreover, psyehophysiologieal signals are eontinuously available and ean be obtained in a non-intrusive manner, pre-requisite for their use in operational environments.

The objeetive of this study is to advanee schemes whieh deteet ehanges in OFS by monitoring psychophysiological signals in real-time. Reviews on similar methods ean be found in, e.g., Wilson and Russell (2003a) and Wilson and Russell (2007). Many of these

methods employ pattern recognition to classify mental workload into one of several discrete categories. For instance, given an experiment with easy, medium and hard tasks, and assuming the tasks induce varying degrees of mental workload on a subject, these methods classify which task is being performed for each epoch of psychophysiological data. The most common classifiers are artificial neural networks (ANN) and multivariate statistical techniques such as stepwise discriminant analysis (SWDA). ANNs have proved especially effective at classifying OFS as they account for the non-linear and higher order relationships often present in EEG/EOG data; they routinely achieve classification accuracy greater than 80%.

However, the discrete output of these classification schemes is not conducive to real-time change detection. They accurately classify OFS, but they do not indicate *when* OFS has changed; the change points remain ambiguous and left to subjective interpretation. Thus, the present study introduces several online algorithms which objectively determine change in OFS via real-time psychophysiological signals.

The following sections describe the dataset evaluated, discuss the statistical properties of psychophysiological signals, and detail the various algorithms which utilize these signals to detect real-time change in OFS. For each algorithm, results are presented regarding their efficacy and a discussion is provided. Finally, the study is concluded with a review and comparison of each method.

2. Data collection, processing, and analysis

The dataset utilized in the following analyses originated from experiments conducted at Wright-Patterson AFB in 2008. Data was available for three subjects. Each subject performed two fourteen-minute trials with unmanned aerial vehicle (UAV) tasks presented at three levels: low, medium, and high, denoted as LL, ML, and HL, respectively. The LL was the baseline state and subsequently encompassed most of each trial. The ML and HL were presented four times each, in a balanced order, lasting approximately 20 seconds each time. Each trial began in the LL and after every ML or HL was presented, it returned to the LL. The experimental design assumed that varying task load induced corresponding levels of cognitive load on the subject. Throughout this report, cognitive load is assumed to be a proxy measure of OFS; thus detecting a change in task load, is synonymous with detecting a change in OFS.

The tasks involved monitoring four UAVs executing a bombing mission. During the ML and HL, subjects performed a visual search of a radar image uploaded at designated waypoints by the UAVs. The subjects panned the radar image, located potential targets, and marked six of them for bombing, according to pre-determined priorities. This had to be accomplished within the 20 second timeframe. The HL was more difficult than the ML as its radar image contained more distracter targets, in addition to imposing more complex targeting priorities. The baseline LL condition simply required the subject to monitor the UAVs flight paths until each waypoint, i.e. ML or HL, was encountered. Overall, the tasks were very visual in nature and were expected to engage the visual processing centers of the brain.

The data collected for each trial consisted of eight physiological channels of EEG, EOG, and electrocardiogram (ECG) recorded at a sampling frequency of 200 Hz. The EEG channels were recorded from five electrodes: F_7, F_z, P_z, T_5, and O_2. These electrodes were affixed to the subject's scalp according to the 10/20 International electrode system

shown in Figure 2.1. Vertical and horizontal EOG data, termed VEOG and HEOG respectively, were collected for two purposes: primarily, as a measure of cognitive load, and secondly, to eliminate blink artifacts in the EEG signals. Finally, one channel of ECG was collected to measure heart rate. Only the EEG and EOG data were used in the following analyses.

From this point forward, each trial will be denoted by a letter followed by a number; the letter represents the subject A, E, or F, and the number identifies whether it was the subject's first or second trial. For instance, A01 denotes data from the first trial of subject A.

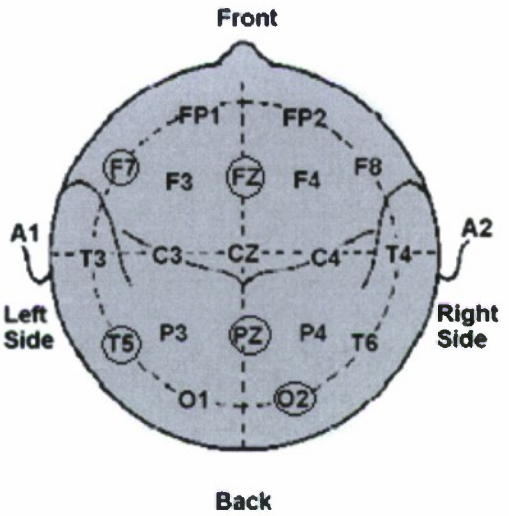

Figure 2.1 EEG electrode diagram

2.1. Preprocessing the data

Blink artifacts contaminate EEG signals when the electrical activity from a subject's eye is recorded by the EEG electrodes affixed to their scalp. This interference greatly distorts the EEG data and negatively affects subsequent analyses. Many filters have been developed to remove blink contamination, of interest here however, are filters that remove the artifacts online, rather than in a post-processing scheme. One form of online artifact removal is adaptive filtering. The adaptive filter incorporates the VEOG and HEOG signals as reference inputs to de-contaminate an EEG signal, $s(i)$, for every time moment i. The artifact-free signal, $e(i)$, results by

$$e(i) = s(i) - \hat{r_v}(i) - \hat{r_h}(i) \tag{2.1}$$

where,

$$\widehat{r_v}(i) = \sum_{m=1}^{M} h_v(m) r_v(n+1-m)$$
$$\widehat{r_h}(i) = \sum_{m=1}^{M} h_h(m) r_h(n+1-m) \quad (2.2)$$

are filtered VEOG and HEOG reference signals, respectively. The $h_v(m)$ and $h_h(m)$ terms represent finite impulse response (FIR) filters of length *M*, which are updated for every time period *i*, to filter the raw VEOG and HEOG signals.

Updating the FIRs is accomplished through a recursive least-squares (RLS) algorithm presented in Table 2.1. The underbars denote column vectors and $R(i)$ is a matrix. In the present analysis, the forgetting factor, λ, was set to .9999 and *M* was set to a length of three. See He, Wilson, and Russell (2007) for complete details.

Figure 2.2 (a) displays an EEG signal recorded at the F_7 electrode, and (b) displays the VEOG signal recorded during the same period of time. Blink contamination is clearly present in (a), where sharp peaks indicate eye blinks. Notice these same blinks are reflected in (b), the VEOG signal, whose intent is to detect eye activity. The adaptive filter uses the VEOG signal to identify the blinks and to remove them from affected EEG signals, like F_7, where blink activity is considered interference. Figure 2.3 depicts the contaminated F_7 signal in (a), but this time with the corresponding filtered F_7 signal in (b), after adaptive filtering. Notice that the sharp peaks characteristic of eye blinks are no longer present.

Once eye blink artifacts were removed, the psychophysiological signals were subjected to Discrete Fourier Transform (DFT) for every epoch, usually three to five seconds. The DFT transforms signals from the time domain to the frequency domain. To accomplish this, the DFT assumes that the time domain signal is a sum of many sinusoids; this assumption is generally deemed appropriate for psychophysiological signals. The DFT and its inverse are computed by

$$X(\Omega) = \sum_{n=-\infty}^{\infty} x[n] e^{-jn\Omega}$$
$$x[n] = \frac{1}{2\pi} \int_0^{2\pi} X(\Omega) e^{jn\Omega} \, d\Omega \quad (2.3)$$

where Ω is the discrete-frequency variable and $x[n]$ is a discrete time series, in this case obtained by sampling a continuous EEG\EOG signal (Phillips et al., 2007).

Utilizing the DFT, a frequency spectrum was created for each epoch where the powers of particular wavebands, e.g. 8-12 Hz, known as *alpha*, *beta*, etc., were computed. Waveband power was computed as the average of each frequency's power, for frequencies falling within the waveband. The powers of particular EEG\EOG wavebands have been shown to correlate with changes in cognitive load (see Table 2.2). For instance, alpha waveband power, 8-12 Hz, increases with relaxation while theta waveband power, 5-8 Hz, generally decreases with relaxation (Smith et al., 2001).

Table 2.1 Adaptive filtering algorithm

1. **Initialize**:
   - 1.1. $\underline{H}(i-1) = 0$ where $\underline{H} = \begin{bmatrix} \underline{H_v} \\ \underline{H_h} \end{bmatrix}$ and $\underline{H_v}$ and $\underline{H_h}$ are vectors of the filter coefficients, $h_v(m)$ and $h_h(m)$, respectively
   - 1.2. $[R(i-1)]^{-1} = \boldsymbol{I}/\sigma$ where $\boldsymbol{I}$ is a $2M$ x $2M$ identity matrix and $\sigma = 0.01$
   - 1.3. M and λ are user-defined
   - 1.4. $i = M$
   - 1.5. n is the length of the signal to be filtered
2. **Calculate** $\underline{K}(i)$:

$$\underline{K}(i) = \frac{[R(i-1)]^{-1}\underline{r}(i)}{\lambda + \underline{r}(i)^T[R(i-1)]^{-1}\underline{r}(i)}$$

where $\underline{r}(i) = \begin{bmatrix} \underline{r_v}(i) \\ \underline{r_h}(i) \end{bmatrix}$ where $\underline{r_v}(i)$ and $\underline{r_h}(i)$ are vectors of the VEOG and HEOG signal, respectively:

$$\underline{r_v}(i) = [r_v(i), r_v(i-1), \ldots, r_v(i+1-M)]^T$$

$$\underline{r_h}(i) = [r_h(i), r_h(i-1), \ldots, r_h(i+1-M)]^T$$

3. **Calculate** $e\left(\frac{i}{i-1}\right)$:

$$e\left(\frac{i}{i-1}\right) = s(i) - \underline{r}(i)^T\underline{H}(i-1)$$

4. **Calculate** $\underline{H}(i)$:

$$\underline{H}(i) = \underline{H}(i-1) + \underline{K}(i)e\left(\frac{i}{i-1}\right)$$

5. **Update** $[R(i)]^{-1}$:

$$[R(i)]^{-1} = \lambda^{-1}[R(i-1)]^{-1} - \lambda^{-1}\underline{K}(i)\underline{r}(i)^T[R(i-1)]^{-1}$$

6. **Calculate** $e(i)$:

$$e(i) = s(i) - \underline{r}(i)^T\underline{H}(i)$$

Note: This is the vector equivalent of (2.1)

7. $i = i + 1$
8. **While** $i < n$ **repeat** steps 2 through 8

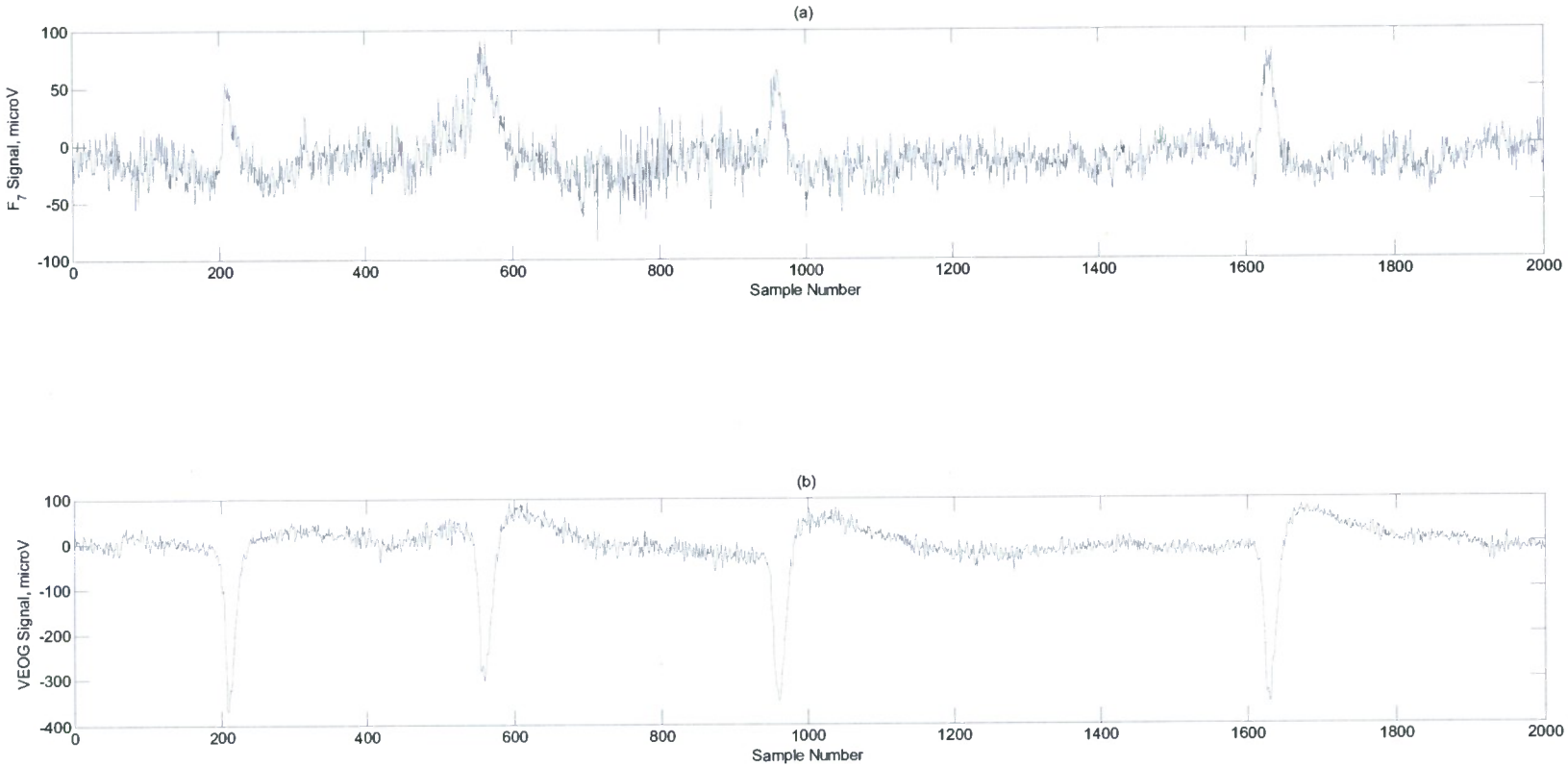

Figure 2.2 (a) F_7 and (b) VEOG signals with blink artifacts

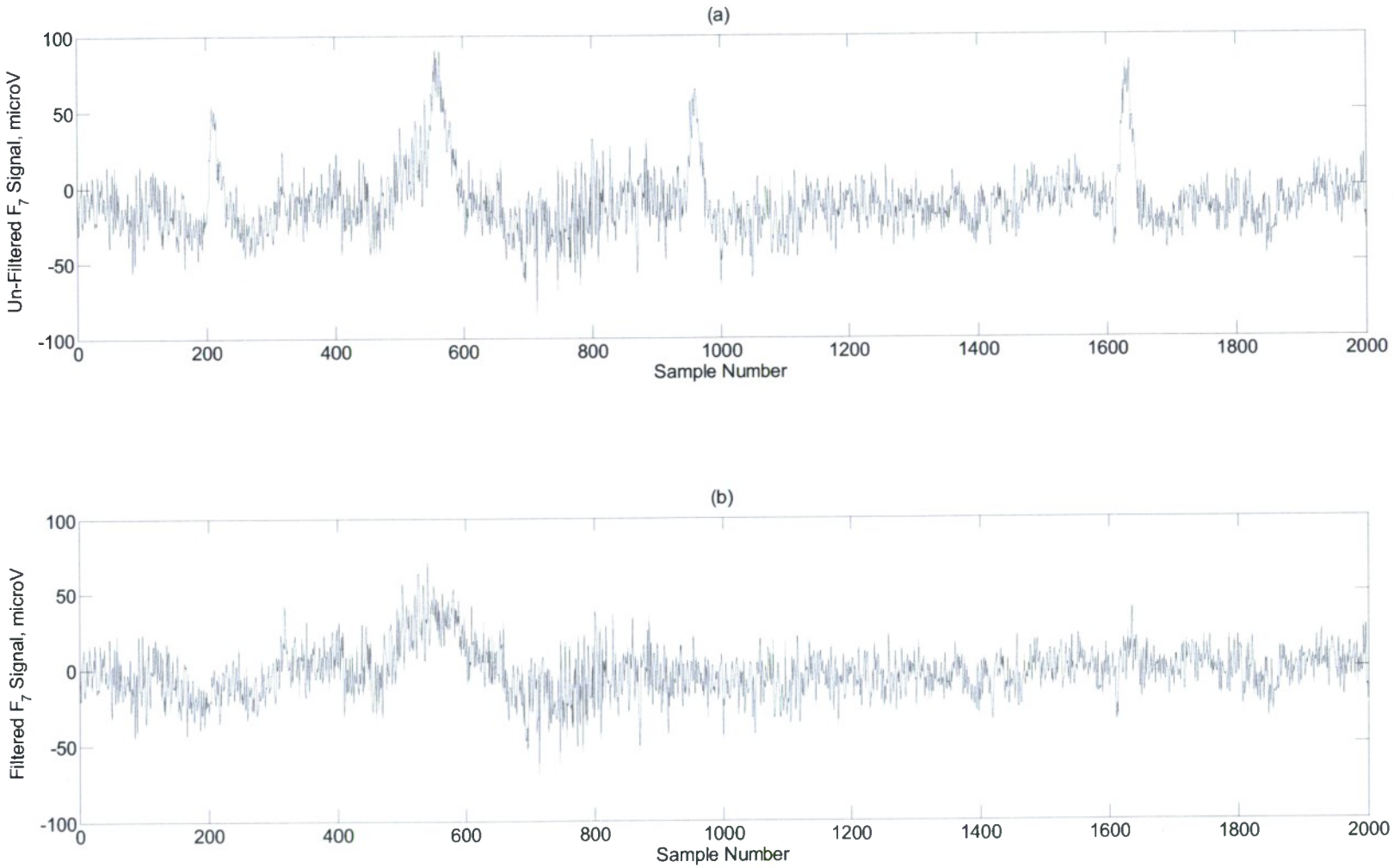

Figure 2.3 (a) Un-filtered and (b) filtered F_7 signals

Table 2.2. EEG wavebands

Waveband	Frequency (Hz)	Interpretation
Delta	1-4	Slow wave sleep
Theta	4-8	Arousal
Alpha	8-12	Relaxation
Beta	12-30	Active, alert, working
Gamma	30+	Cognitive and/or motor function

From this point forward, the waveband powers computed from psychophysiological signals recorded at particular electrodes are referred to as *features*. It is possible to compute hundreds of features for each epoch, as there are many different combinations of wavebands and electrodes. Methods which measure and detect real-time changes in OFS must determine which of these features are indicators of cognitive load, and which are noise.

2.3. Statistical Analysis of the Post-Processed Data

Empirical analysis was conducted on several psychophysiological features to characterize their properties and behavior with respect to subject and task load. For this analysis, the features were computed in three-second epochs. It is common to find one-second epochs in the literature, however, they often exhibit highly variable and erratic behavior. Moreover, short epochs induce autocorrelation in the features. In contrast, longer epochs are more stable, but if they are too long, they will be infeasible in operational settings where change occurs instantaneously. Three-second epochs were found to achieve the best balance between short and long epochs.

The features chosen for this analysis are traditional features used in OFS classification methods. The wavebands of these features are: delta, 2-4 Hz, theta, 5-8 Hz, alpha, 9-13 Hz, beta, 14-32 Hz, and gamma, 33-43 Hz. Features from these wavebands have shown to correlate strongly with cognitive load, thus they are often used to classify OFS.

Figure 2.4 displays the distributions of theta, alpha, and beta waveband powers computed across all electrodes for every epoch. The data from every trial were combined to form the distributions, after the data were standardized to zero mean and unit variance with respect to each trial. As shown, the distributions are non-normal, displaying a strong positive skew. The beta waveband power is especially skewed and nearly resembles a Poisson distribution. These distributions were subjected to statistical tests for non-normality and each was strongly significant at the .05 level, indicating that features from these wavebands originate from non-normal distributions.

Recall that all seven electrodes record the same five wavebands for every epoch. The reason for this seemingly "redundant" data is that brain waves behave differently

when emitted from different regions of the brain. For instance, alpha power is greatest in parietal regions, e.g. P_z, whereas theta power dominates the frontal region, e.g. F_z. While differences exist between topographical locations, there still remains significant correlation, or dependency, between identical wavebands collected at different electrodes; hence not every feature records unique information.

To illustrate the correlation between features, Table 2.3 contains the correlation coefficients of features recorded at VEOG of F01. Notice that every feature was positively correlated with an average correlation coefficient of .819. In this sense, one feature from F01's VEOG electrode could virtually represent all the information recorded there. The understanding of this phenomenon implores selecting a variety of features, from different electrodes, so they each contain unique information about the subject's cognitive state.

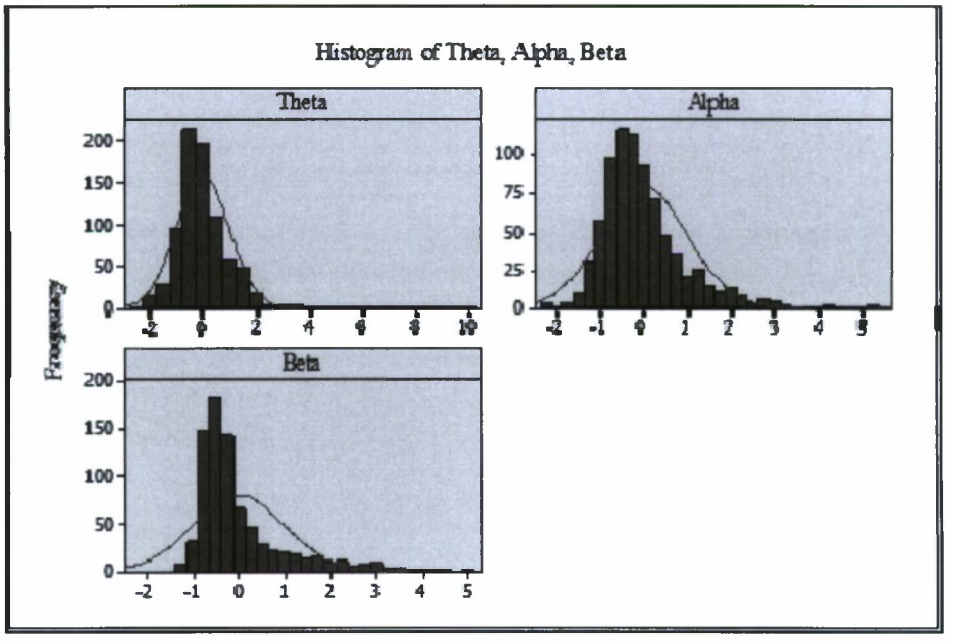

Figure 2.4 Histograms of theta, alpha, and beta wavebands collapsed across all subjects

Many of the features exhibited another form of correlation, called autocorrelation. Autocorrelation occurs when an observation, , recorded at time $i+k$, is dependent upon a previous observation, , where k is the time lag. Thus, if an autocorrelated feature is observed at time i, it will contain information about its value k time moments in the future (Montgomery et al., 2008). Autocorrelation is identified by evaluating the sample autocorrelation function for a finite time series, , defined by

— (2.4)

where c_k and c_0 are computed from

$$c_k = \frac{1}{n}\sum_{i=1}^{n-k}(x_i - \overline{x})(x_{i+k} - \overline{x}) \; for \; k = 0,1,2,\ldots,K \tag{2.5}$$

Beta and gamma features consistently exhibited autocorrelation. Figure 2.5 displays the sample autocorrelation function for a beta feature measured at T_5 of E01. At lag 1, the autocorrelation is the strongest and then exponentially decreases before oscillating around 0. This behavior is characteristic of a first-order, autoregressive time series (AR[1]).

A survey of feature autocorrelation was conducted across subjects. The results are summarized in Table 2.4. For each waveband, a percentage is displayed representing the proportion of electrodes where the feature exhibited significant autocorrelation. For instance, the alpha waveband for A01 generated features which were autocorrelated in three of the seven electrodes, i.e., 42.86%. As shown, the wavebands which generated the least autocorrelated features across subjects were theta and alpha. In contrast, beta and gamma wavebands gave rise to features that were autocorrelated regardless of subject or electrode.

Table 2.3 Inter-node correlation measured at VEOG of F01

	Delta	**Theta**	**Alpha**	**Beta**	**Gamma**
Delta	1.000	0.922	0.939	0.852	0.582
Theta	0.922	1.000	0.949	0.864	0.655
Alpha	0.939	0.949	1.000	0.914	0.663
Beta	0.852	0.864	0.914	1.000	0.853
Gamma	0.582	0.655	0.663	0.853	1.000

It is also important to characterize how the properties of psychophysiological features change with time. More specifically, it is necessary to investigate if these features exhibit stationary behavior. Stationarity implies a statistical equilibrium, where the properties of a time series, such as its autocorrelation or probability distribution, are stable over time. A time series is strictly stationary when the joint probability distribution of observations $x_i, x_{i+1}, \ldots, x_{i+n}$ is the same as the joint distribution of $x_{i+k}, x_{i+k+1}, \ldots, x_{i+k+n}$. However, it is often sufficient to classify a time series as stationary as long as it varies around a fixed mean (Montgomery et al., 2008).

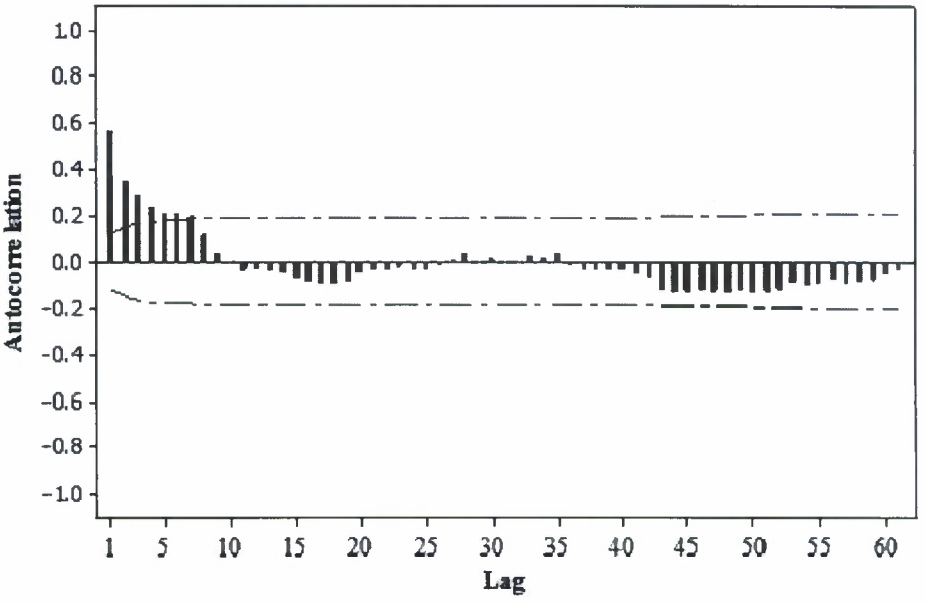

Figure 2.5 Sample autocorrelation function for T_5 beta of E01

Note: the red lines denote .05 significance levels

Table 2.4 Results of autocorrelation survey

	A01	E01	F01	Average
Delta	0.00%	57.14%	71.43%	42.86%
Theta	14.29%	28.57%	42.86%	28.57%
Alpha	42.86%	28.57%	42.86%	38.10%
Beta	100.00%	100.00%	85.71%	95.24%
Gamma	100.00%	100.00%	100.00%	100.00%

By the latter definition, non-stationary features will exhibit some type of trend, for instance, their mean might increase with time. These trends can be determined by plotting a feature's autocorrelation function. If the plot displays slowly-decreasing autocorrelation with increasing lag, a feature is classified as non-stationary. When this analysis was conducted on a sample of features from each waveband, none of the features exhibited the symptoms of non-stationarity. Instead, they each varied around a fixed mean as in the example shown in Figure 2.6.

Finally, an empirical analysis was conducted to characterize the relationship between features and varying task load; these results are in Appendix A. Although no casual relationship between psychophysiological features and task load are firmly established, many studies have asserted strong associations, such as increasing alpha

power with decreasing task load. The features that are most valuable to OFS classifiers, are those that monotonically increase or decrease with changing task load. If this relationship is present, then these features can provide information on a subject's cognitive load. Moreover, it would be ideal if particular features had the same correlation to task load regardless of the subject; however, the results of this empirical analysis found no such "universal" feature. Instead, the features of each subject exhibited unique behavior with respect to task load.

Despite the inability of features to generalize across subjects, there existed at least one feature for each subject which monotonically increased or decreased with task load. Thus by selecting features tailored to each subject, it is possible to gain reliable information on OFS. In fact, nearly all OFS classifiers adhere to this "subject-specific" approach (Wilson, Russell, 2003a; Smith et al., 2001).

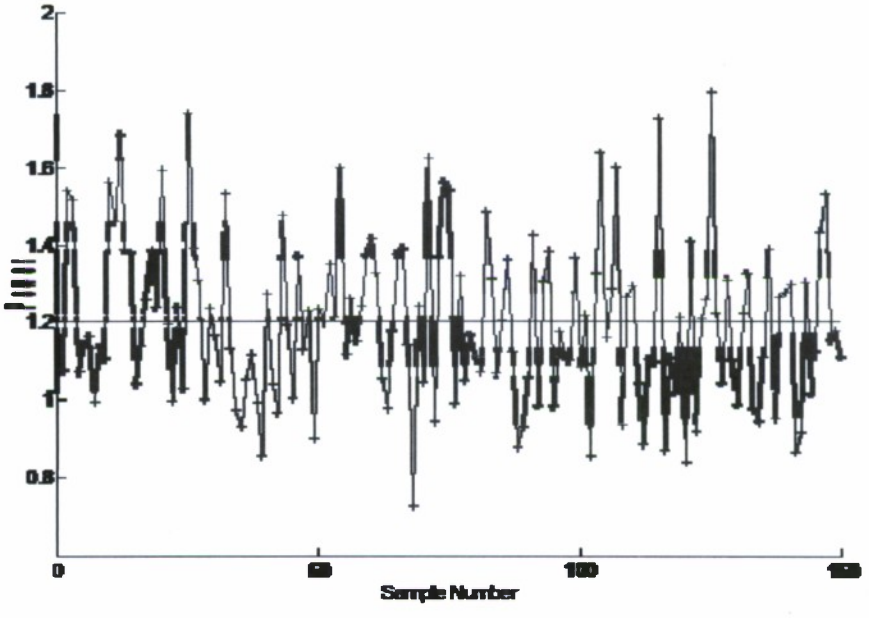

Figure 2.6 Stationary plot of F_z theta for A01

3. Independent Component Analysis of Multidimensional Psychophysiological Time Series

Independent component analysis (ICA) is a method used to extract the underlying components of signals. ICA assumes that physical processes, like the brain, are comprised of distinct operators which emit signals independent of each other. When these signals are recorded by sensors, they become "mixed" and indistinguishable. ICA is a method for separating these mixed signals into their underlying components (Stone, 2002).

For instance, two EEG electrodes record brain activity at different locations on the scalp. The signals recorded from these electrodes, denoted as $x_1(i)$ and $x_2(i)$, are assumed to be a mixture of components, $s_1(i)$ and $s_2(i)$, originating from independent operators in the brain. $x_1(i)$ and $x_2(i)$ are expressed as linear combinations of the source components,

$$\begin{aligned} x_1(i) &= a_{11}s_1(i) + a_{12}s_2(i) \\ x_2(i) &= a_{21}s_1(i) + a_{22}s_2(i) \end{aligned} \tag{3.6}$$

The objective of ICA is to reveal the source components using only the recorded EEG signals, $x_1(i)$ and $x_2(i)$. As shown in (3.6), this requires determining the coefficients a_{ij}. This is a difficult task, however, it is made feasible by assuming the source components are independent of each other, hence "independent components" (Hyvärinen, Oja, 2000).

ICA is more generally defined by considering the system of equations in (3.6) in vector-matrix notation,

$$\boldsymbol{x}(i) = \boldsymbol{A}\boldsymbol{s}(i) \tag{3.7}$$

where $\boldsymbol{A}$ is the mixing matrix, the recorded psychophysiological signals are in vector $\boldsymbol{x}(i)$, and the independent components are in vector $\boldsymbol{s}(i)$, for each epoch i. Once $\boldsymbol{A}$ is determined, its inverse, $\boldsymbol{W}$, is used to compute the independent components for every realization, $\boldsymbol{x}(i)$, by

$$\boldsymbol{s}(i) = \boldsymbol{W}\boldsymbol{x}(i) \tag{3.8}$$

Notice that $\boldsymbol{x}(i)$ is what is observed, but $\boldsymbol{s}(i)$ is what needs to be computed. Thus, similar to PCA, the independent components are determined indirectly and called latent variables (Hyvärinen, Oja, 2000). Moreover, ICA is the most common means of blind source separation (BSS), where underlying factors are determined "blindly". An assumption of BSS, which ICA necessarily makes, is that the number of independent components is limited to the number of recorded signals. For instance, in the present data there were seven EEG and EOG signals recorded, thus, ICA assumes each were a mixture of seven independent components (Stone, 2002).

Some fundamental assumptions of ICA have been mentioned, but more assumptions are necessary to make ICA feasible. In order to achieve components which are independent, it is necessary to assume the components are non-Gaussian distributed (i.e. non-normal). In this study, independent components are established by constructing a mixing matrix which maximizes non-Gaussianity. Next, the source signals are assumed to propagate through a medium, in this case brain tissue, instantaneously, before being *linearly* mixed at the electrodes. Finally, the source signals are assumed to be stationary.

ICA is made robust to moderate violations of these assumptions through the FastICA algorithm, detailed below. The major assumption which cannot be compromised, however, is the number of source signals assumed to exist. For the present

purpose, this number is less important than finding components which are meaningful to OFS change detection (Makeig et al., 1996; Vigário et al., 2000).

Finally, ICA differs fundamentally from PCA, despite their similarities. They are both multivariate analyses, whose latent variables are linear combinations of observed variables. However, PCA's primary goal is dimensionality reduction, whereas ICA creates a generative model to reveal the underlying factors of a process. Additionally, PCA merely de-correlates its latent variables. Under the assumption of normality, de-correlation is sufficient to achieve independence. However, psychophysiological signals have consistently displayed non-normal behavior, implying that PCA's de-correlation does *not* result in truly independent components. In contrast, ICA imposes stricter constraints which achieve un-rotatable, de-correlated, and *independent* components (Stone, 2002).

Pre-Processing the Data for Independent Component Analysis Before conducting ICA, the data must be pre-processed. The data are made to be zero mean, resulting in independent components which are also zero mean; this is done to facilitate the estimation of $\boldsymbol{A}$. After it is estimated, the mean is added back to the independent components.

Next, the data are whitened to make $\boldsymbol{A}$ orthogonal which significantly reduces the number of parameters ICA must estimate. Whitening linearly transforms $\boldsymbol{x}(i)$ into a new vector, $\tilde{\boldsymbol{x}}(i)$, that is uncorrelated and has unit variance. Whitening can be done through eigenvalue decomposition (EVD). Using the same notation, $\tilde{\boldsymbol{x}}(i)$ is computed by

$$\tilde{\boldsymbol{x}}(i) = \boldsymbol{U}\boldsymbol{L}^{-1/2}\boldsymbol{U}'\boldsymbol{x}(i) \tag{3.9}$$

This step can also be used to dimensionally reduce the independent components, which serves to decrease noise and prevent overfitting (Hyvärinen, Oja, 2000).

The FastICA Algorithm Once the data are pre-processed, an efficient algorithm called FastICA is used to estimate $\boldsymbol{W}$. FastICA determines the columns of $\boldsymbol{W}$, denoted as $\boldsymbol{w}_j$, one-by-one, by maximizing the non-Gaussiantity of the projection, $\boldsymbol{w}_j'\boldsymbol{x}(i)$. There are several measures of non-Gaussianity, among the most common is negentropy. Negentropy is an entropic metric that captures the "randomness" of a variable. Because Gaussian variables are the most random, they have the highest entropy among all random variables. Therefore, FastICA indirectly maximizes non-Gaussianity by pursuing the least entropic variables as measured by negentropy. As mentioned, the non-Gaussianity of components is essential to achieve their independence. An approximation of negentropy, $J(y)$, is computed by

$$J(y) = (E[G(y)] - E[G(v)])^2 \tag{3.10}$$

where G is a non-quadratic function, y and v are both zero mean and unit variance variables, and v is Gaussian distributed. In this metric, if y is also Gaussian then negentropy is zero; any other distribution of y results in positive negentropy, with magnitude proportional to y's deviation from Gaussianity. In the present project, the non-

quadratic function was specified as, $G(y) = y^4$, rendering (3.10) a kurtosis-based approximation of negentropy.

The FastICA algorithm is presented in Table 3.2. Steps 2.3 through 2.5 iteratively refine a randomly chosen vector, $\boldsymbol{w}_j$, until it converges. Convergence occurs when a new $\boldsymbol{w}_j$ points in the same direction as the vector from the previous cycle of steps 2.3 and 2.4. Once convergence occurs, step 2.8 employs a deflation procedure to prevent different $\boldsymbol{w}_j$ vectors from reaching the same maximum; deflation de-correlates the outputs of $\boldsymbol{w}_1'\boldsymbol{x}(i), \boldsymbol{w}_2'\boldsymbol{x}(i), \ldots, \boldsymbol{w}_n{}'\boldsymbol{x}(i)$. Finally, step 2.9 renormalizes $\boldsymbol{w}_j$ after deflation. The algorithm iterates until $\boldsymbol{W}$ is fully defined.

Independent Component Analysis to Facilitate Real-Time OFS Change Detection

In past research, ICA has been conducted on physiological signals for two purposes: to identify and extract artifacts, such as eye blinks, and to facilitate the analysis of event-related potentials (ERP). There are many artifacts in psychophysiological signals, such as neck muscle activity, eye blinks, heart rate, and line noise originating from recording equipment. In some cases, the artifacts' amplitude exceeds the brain activity, the very thing intended to be measured. Therefore, it is critical to identify and extract these artifacts before using psychophysiological features to detect change in OFS. ICA can automatically extracts most of the prominent artifacts, producing less complex and less noisy source signals (Vigário et al., 2000).

ICA is also commonly used to analyze ERPs in EEG signals. ERPs are spikes in brain activity that result from the onset of a stimulating event, like hearing a "beep". It is believed that ERPs originate from different sensory systems in the brain, corresponding to the sense being stimulated, like sight, sound, and touch. To identify the location of these sensory systems, experiments are conducted which simultaneously stimulate multiple senses of a subject while recording their EEG. The ERPs are contained in the EEG, but are mixed upon recording. ICA is then applied to "un-mix" these signals and generate independent components corresponding to the ERPs which originated from each sensory system (Vigário et al., 2000; Makeig et al., 1996).

To the best of our knowledge, ICA has yet to be employed to facilitate the online change detection of OFS. Nevertheless, ICA is a promising candidate for this application since it has been effective at artifact extraction and generating independent components that correspond to external stimuli. The objective is to use ICA to extract source signals which are free of artifacts and responsive to varying task load.

Table 3.1 FastICA algorithm

1. **Initialize:**
   1.1. $j=1$
   1.2. n is the number of independent components

2. **Run Algorithm:**
   2.1. **for** $j < n$
   2.2. Choose a random vector $\boldsymbol{w}_j$
   2.3.
   $$\boldsymbol{w}_j{}^{+} = E\left[\boldsymbol{x}(i)G'\left(\boldsymbol{w}_j{}'\boldsymbol{x}(i)\right)\right] - E\left[G''\left(\boldsymbol{w}_j{}'\boldsymbol{x}(i)\right)\right]\boldsymbol{w}_j$$
   Note: expectations are estimated by sample means.
   2.4.
   $$\boldsymbol{w}_j = \frac{\boldsymbol{w}_j{}^{+}}{\left\|\boldsymbol{w}_j{}^{+}\right\|}$$
   2.5. **if** not converged, go back to 2.3
   2.6. **else if** $j==1$ go back to 2.1
   2.7. **end if/else**
   2.8.
   $$\boldsymbol{w}_j = \boldsymbol{w}_j - \sum_{k=1}^{j-1} \boldsymbol{w}_j{}'\boldsymbol{w}_k\boldsymbol{w}_k$$
   2.9.
   $$\boldsymbol{w}_j = \frac{\boldsymbol{w}_j}{\sqrt{\boldsymbol{w}_j{}'\boldsymbol{w}_j}}$$
   2.10. $j=j+1$
   2.11. **end for**

3. **Output:** the inverse mixing matrix, $\boldsymbol{W}$

Analyzing the Independent Components For each subject, all seven EEG and EOG signals were submitted to ICA and the resulting independent components were analyzed. In every case, at least two independent components correlated directly with eye activity measured at VEOG and HEOG; this is consistent with the visual demands of the UAV tasks. This phenomenon repeated itself when only EEG signals were submitted to the ICA, hence providing strong evidence that eye activity contaminated the EEG data. Figure 3.1 depicts two independent components extracted from the five EEG signals of E01; it clearly displays their correlation to the VEOG and HEOG signals that were not included in the analysis.

The independent components that were highly correlated with VEOG and HEOG signals were deemed "artifact" signals and were excluded from further analysis. The remaining components were empirically analyzed to determine what brain process they represented. This was accomplished by constructing a frequency spectrum for each component, to reveal the frequencies that dominated them. Again, the DFT, as shown in (2.3), was used to convert the components from the time domain to the frequency domain.

Figure 3.2 depicts the frequency spectrums of four independent components from A01. Each component was partitioned in HL, ML, and LL sets and submitted to the DFT in three-second epochs. The results from the DFT were averaged across all epochs to form three frequency spectrums, one for each task load. As shown, each independent component exhibited unique frequency spectrums, some of which varied by task load. For instance, plot (c) displays the strong influence 10-12 Hz waves had on the third independent component. Moreover, the power of these waves varied by task load.

In fact, every subject had at least one component that was strongly influenced by 8-12 Hz waves, corresponding to the alpha waveband. Recall that this waveband correlates to relaxation; as a subject becomes more relaxed, their alpha power increases. This is consistent with plot (c) of Figure 3.2, where alpha power is highest in the LL and lowest in the HL.

Finding alpha's presence in the psychophysiological signals of cognitive tasks is not surprising, but revealing its prominence with such clarity is rare. Alpha power is frequently obscured by noise when signals are mixed and recorded at the scalp. ICA segregated these noisy signals from the source signals and clearly displayed alpha's contribution. Makeig et al. confirmed this result, reporting that ICA revealed "alpha activity (near 10 Hz) not obvious in the EEG data" (1996). This highlights another result, that alpha power most often peaked around 10 Hz in the independent components. This particular frequency's significance has been established in other research, where it has been shown to correlate with performance on cognitive tasks (Makeig, Inlow, 1993). Research has also shown 10 Hz to be associated with spatial tasks, more so than verbal ones (Gevins, 1997). This is consistent with the present results, since the UAV tasks were largely spatial in nature.

Analyzing the frequency spectrum of a subject's independent components may prove useful in detecting changes of OFS. An abundance of research has employed the frequency domain of EEG signals for classification, most commonly to analyze the traditional wavebands of delta, theta, alpha, beta and gamma. Perhaps a more effective strategy would be to use independent components to derive subject-specific and task-specific wavebands. An analysis similar to the one depicted in Figure 3.2 could be done

for each subject to identify frequency bands which are sensitive to changes in task load. Wilson and Fisher report using subject-specific wavebands derived from latent variables to increase classification accuracy by 29% over traditional wavebands (1995). In another case, estimating the error rates on task performance was improved significantly by using a subject's entire EEG spectrum over predefined, narrow wavebands (Jung et al., 1997).

However, one shortcoming of this method is the subjective process of selecting frequency bands from the independent components. It requires expert knowledge to identify which components represent artifacts, to understand how brain activity responds to different tasks, and finally, to decide which frequency bands are sensitive to task load. These ambiguities can be mitigated through dimensionality reduction, where multiple psychophysiological signals are reduced to result in a pair of independent components. From this reduction, an independent component may emerge with properties which are easily identifiable and robust across subjects. The next section discusses this approach.

Dimensionally Reduced Independent Components During the pre-processing for ICA, eigenvalue decomposition (EVD) was used to whiten the data. EVD is equivalent to PCA, and like PCA, it can dimensionally reduce the psychophysiological signals to a subset of important components. Dimensionality reduction prevents overfitting of the independent components and decreases noise, not to mention, reduces the vast number of features into a small, yet crucial subspace. Most importantly, dimensionality reduction reveals the independent components responsive to task load with less ambiguity.

EVD was used to reduce the seven EEG and EOG signals to a pair of signals for submission to ICA. Reducing the signals from seven to two was justified by a Pareto chart analysis which, for every subject, displayed a clear break at two when plotting the eigenvalues in descending order. The resulting pair of signals was submitted to ICA which generated two independent components. These components were then plotted in and empirically assessed for any correlation they had with task load.

This method yielded consistent results when conducted on each trial. In every case, one independent component exhibited clear variation with changing task load, while the other was mostly noise. Figure 3.3 displays the results of A01, where (a) is the component responsive to task load and (b) is the noise component. Notice that the independent component in (a) became attenuated during the HL and ML, compared to its variation in the LL. The second component in (b) exhibited similar behavior, but was not consistent for every task and was obscured by noise. The utility of the first component for real-time OFS change detection is promising since: one exists for each subject, its behavior is distinguishable by task load, and finally, the ICA method is automatic and leaves little error-prone subjectivity.

Further investigation of the dimensionally reduced independent components disclosed strong correlations with the VEOG and HEOG signals. Previously, the independent components correlated with VEOG and HEOG were excluded from further analysis since they were deemed "artifacts". However, in the present analysis these components were extracted as the source signals for all the EEG\EOG data, thus they were retained.

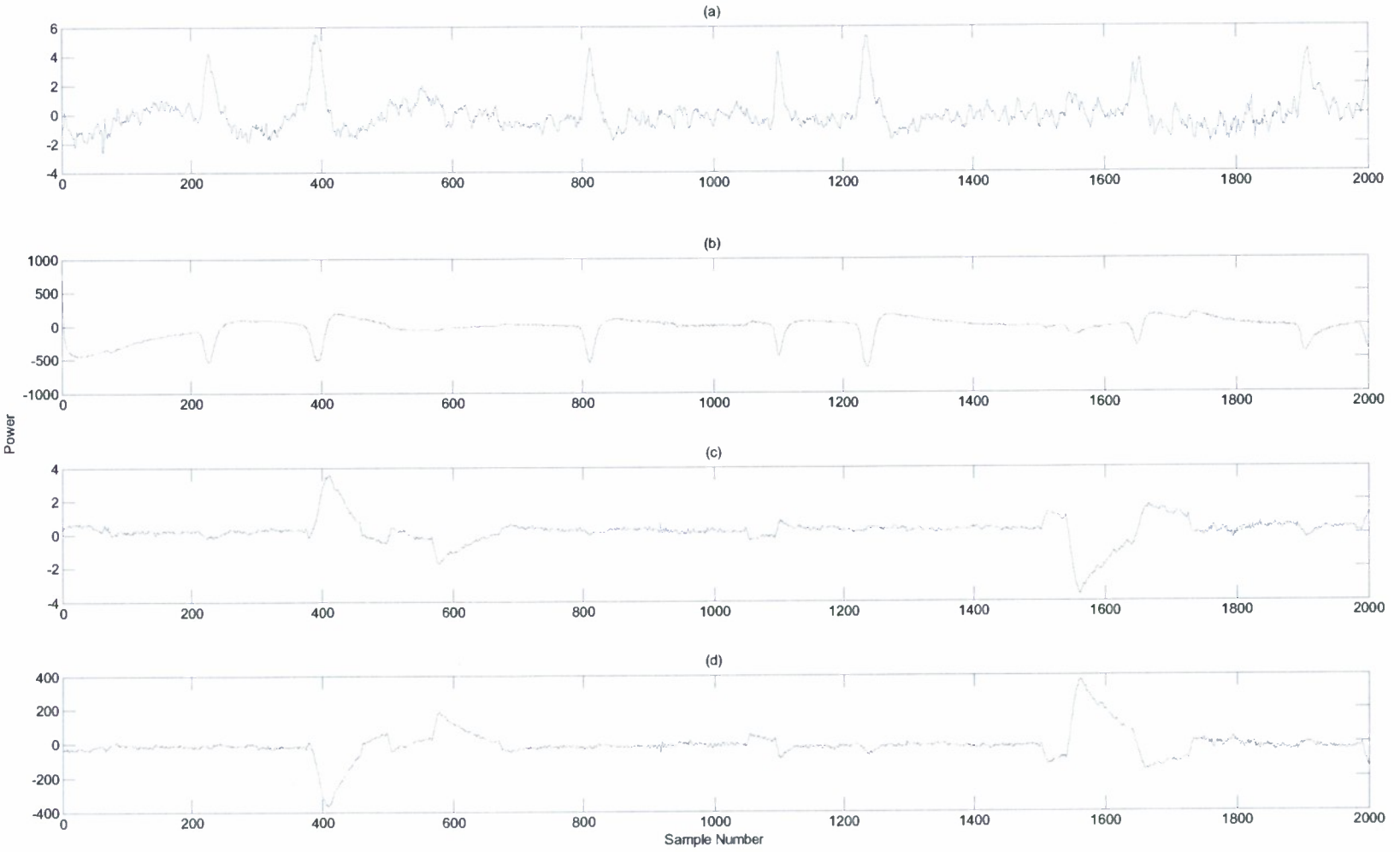

Figure 3.1 (a) IC1 (b) VEOG (c) IC2 and (d) HEOG signals of E01

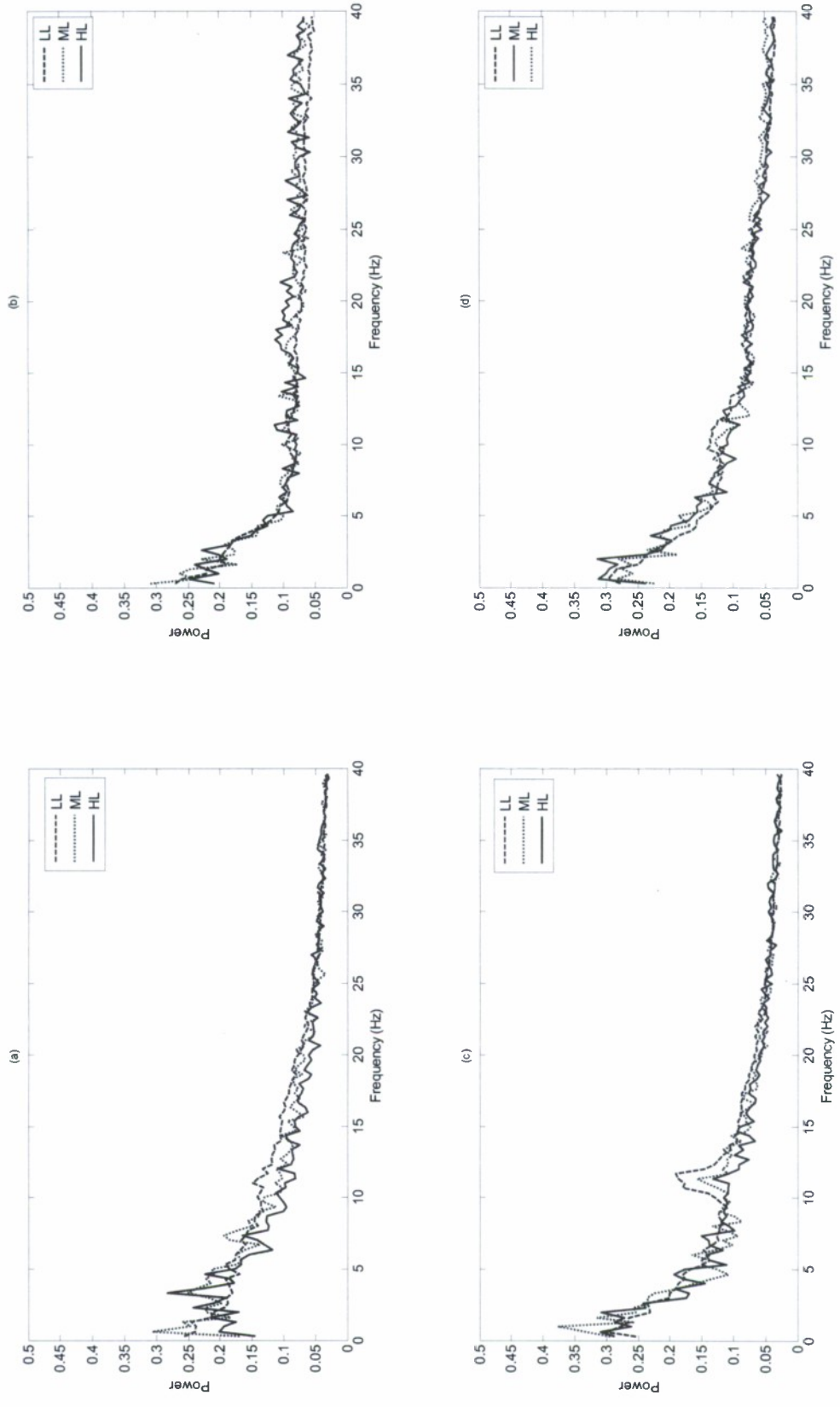

Figure 3.2 Frequency spectrum of independent components from A01, (a) IC1 (b) IC2 (c) IC3 (d) IC4

Table 3.3 and Table 3.4 contain the correlations between the original psychophysiological signals and the first and second independent components, respectively. The correlations are averaged across trails by subject. Subject F is considered separately since, when his components were plotted, they exhibited unique behavior. Subject F has consistently been the most difficult to accommodate in a varicty of methods, and the correlation analysis provided clues as to why.

Table 3.3 Correlation analysis of dimensionally reduced first independent componcnt

	VEOG	HEOG	F_z	F_7	P_z	T_5	O_2
Subjects A, E	0.903	0.107	0.835	0.385	0.389	0.297	0.179
Subject F	0.977	0.075	-0.380	-0.137	-0.261	-0.228	-0.174

Table 3.4 Correlation analysis of dimensionally reduced second independent component

	VEOG	HEOG	F_z	F_7	P_z	T_5	O_2
Subjects A, E	0.373	0.992	0.067	0.860	0.281	0.036	0.475
Subject F	0.213	-0.991	-0.104	-0.870	0.045	-0.159	0.227

In Table 3.3, notice the strong correlation the first indepcndent component had with the VEOG signal across subjects. Another notable correlation for subjects A and E, is with the EEG signal recorded at F_z. Howcver, this is not true for subject F, who averaged less than half the correlation at F_z than subjects A and E.

The second independent component, in Table 3.4, correlated with one EOG and one EEG signal, similar to the first component. But in this case, the correlation was with HEOG and F_7, respectively. Noticc that subject F's bchavior was not apprcciably different than the other subjects.

This analysis elucidates scveral important findings, principally, that thc sourcc signals of the EEG\EOG were strongly correlated to eye activity. This is useful knowledge, as otherwise, VEOG and HEOG contaminated signals may be discarded as artifacts rather than treated as important. As mentioncd, one of the two independcnt components always varied with task load and the other was mostly noise. For subjects A and E, the independent component which best varied with task load was the first. Thc correlations in Table 3.3 indicate that this component was a mixture of signals from VEOG and F_z. Recall that the F_z electrode mostly records theta waves which arc sensitivc to changes in cognitive load. This was true except for subject F, who did not have a strong F_z presence in the first component. In fact, subject F's first component was noisy and it was his second component, the one dominated by HEOG and F_7, that best varied with task load. Two things can be inferred from this: first, ICA can rcveal the source of between-subject differences, and second, those differences can be overcome with at least

one independent component that is sensitive to task load. Henceforth, the independent component that best varied with task load, post-dimensionality reduction, is referred to as the TVIC (task-varying independent component).

4. Real-Time Metrics for Psychophysiological Signals and Algorithms for Temporal Change Detection in Time Series Data

The Peak Detection Method Now that a robust signal, the TVIC, can be identified for each subject, a quantitative metric is necessary to characterize the signal for each epoch; this metric will be used to facilitate the online change detection of OFS. The TVIC plotted in (a) of Figure 3.3 suggests that a metric of variance might best characterize the signal's response to task load, since variance noticeably increased during the LL and drastically reduced during the ML and HL. Ideally, monitoring a metric like variance, computed for each epoch, would result in a signal with less noise and better discrimination between task loads than the raw TVIC. In essence, the magnitude of the metric would "peak" when the tasks are incurred, thus signifying changes to higher cognitive loads.

Although variance is an obvious choice, there are other quantitative metrics to characterize a signal in real-time, such as the third moment (skewness), the fourth moment (kurtosis), and finally, entropy based metrics. These metrics may also prove sensitive to the onset of tasks and even result in superior change detection.

An objective method was needed to evaluate and compare each metric's ability to produce consistent task-induced peaks. One such technique is called the *peak detection method* (PDM). The PDM compares a measure x_i at epoch i with two thresholds, an absolute threshold defined by parameter γ and a relative threshold defined by parameter δ; the comparisons are made to determine if a peak has occurred.

For the PDM, each metric was computed by a sliding window so that irregularities were smoothed and peaks were induced during the tasks. The PDM was used to objectively evaluate which metric most consistently peaked for all 8 HL and ML tasks, while generating the fewest false alarms. Hence, the results of the PDM indicated which metric best detected change in OFS. The algorithm for the PDM is presented in Table 4.1 (Yu, 2009).

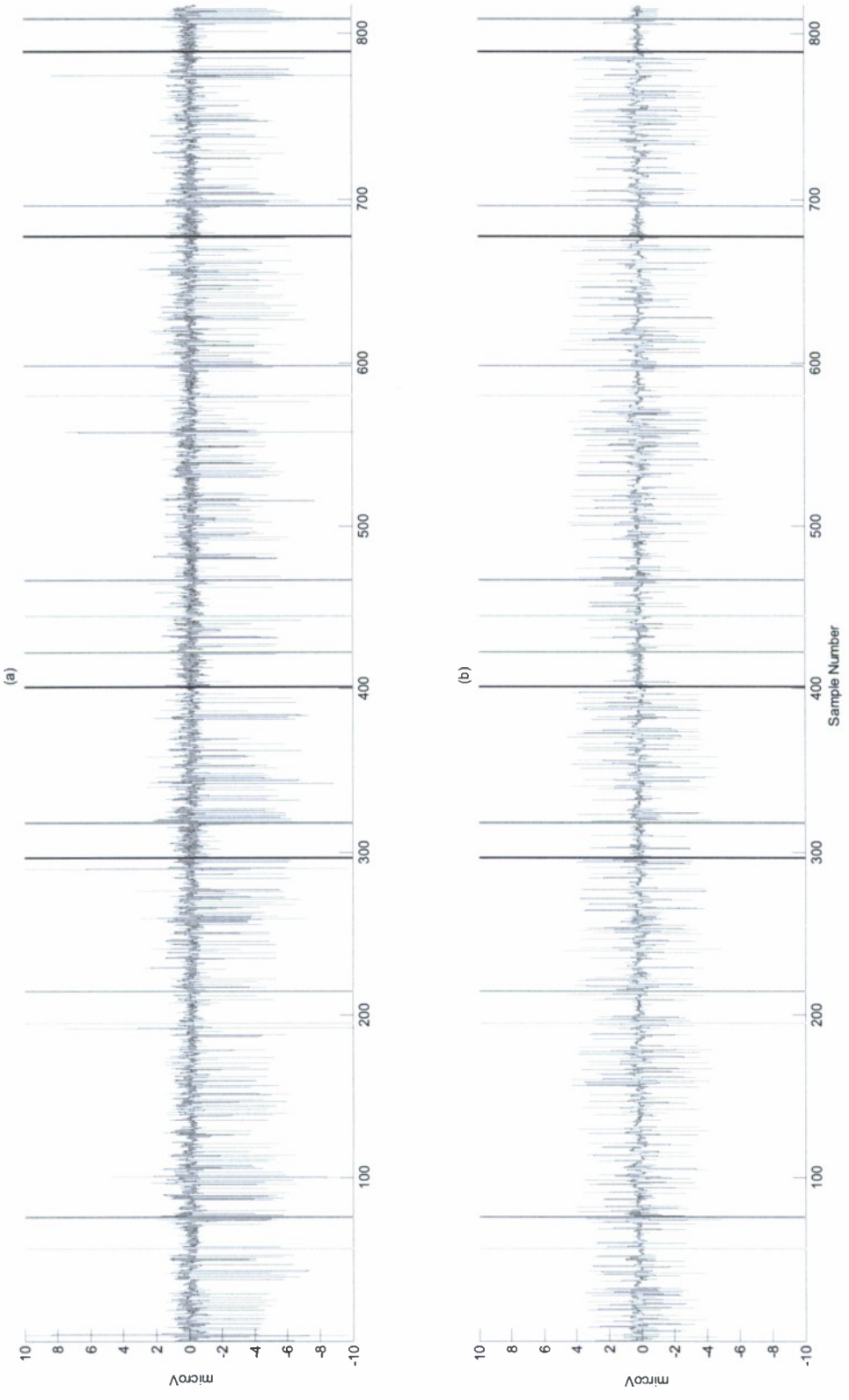

Figure 3.3 Dimensionally reduced independent components of A01, (a) IC1 (b) IC2

Review of Metrics for Characterizing Psychophysiological Signals The simplest metric for characterizing the TVIC is the standard deviation. The standard deviation is the positive square root of the signal's second moment, defined as

$$\sigma = \sqrt{M_2} \tag{4.1}$$

where M_j is

$$M_j = \frac{\sum_{i=1}^{n}(x_i - \mu)^j}{n} \tag{4.2}$$

μ is the mean of the signal, generally estimated by the sample mean, $\bar{x}$, computed from a sample of size n. For the standard deviation calculation, n is traditionally reduced by one in the denominator of (4.2) (Montgomery, 2009).

A related metric is the third moment of the TVIC, termed skewness when normalized, which measures the asymmetric nature of a distribution. The TVIC in plot (a) of Figure 3.3 displays clear skewness in the distribution of its amplitudes during the LL, but much less skewness during the tasks. Skewness is defined as

$$\gamma = \frac{\mathrm{M}_3}{M_2^{3/2}} \tag{4.3}$$

(Montgomery, 2009).

The fourth moment may also prove a valuable metric as it characterizes the shape of a distribution, discriminating distributions that are tall and skinny from those that are short and stout. When standardized, the fourth moment is termed kurtosis, and because it is fourth order, it is always positive. Kurtosis is defined as

$$k = \frac{M_4}{M_2^2} \tag{4.4}$$

(Montgomery, 2009).

The next two metrics are both measures of entropy, where entropy quantifies the "randomness" of the TVIC. The more random the data, the higher entropy it will have, conversely, the more ordered the data, the less entropy it will have; data in perfect order has zero entropy. In recent research, entropy measures of psychophysiological data have been found to correlate with changing states of vigilance, in particular, entropy increased with increasing vigilance (Bruzzo et al., 2008; Zhang et al., 2009). This result suggests entropy metrics may be able to detect changes in OFS.

The first entropy metric considered, called sample entropy (SampEn), is a non-linear metric that measures the regularity of a time series (Richman et al., 2004). SampEn was chosen because it is computationally efficient and can be applied to short, noisy data typical of the TVIC. In addition, SampEn is an improvement over the more traditional

time series entropy, approximate entropy (ApEn); SampEn exhibits less bias towards lower entropy, improves the measure's relative consistency between datasets, and is more robust with regards to varying record lengths (Richman, Moorman, 2000).

SampEn measures the regularity of a time series by computing the conditional probability that two arbitrarily similar epochs of size m, remain arbitrarily similar for the next point in the series. "Arbitrary" is used here to emphasize a degree of error allowed when classifying two epochs as similar. This is controlled by the parameter r, typically a factor between .1 and .25 of the record's standard deviation. By SampEn's definition, a perfectly ordered series will have a conditional probability of one, corresponding to entropy of zero. Time series with less order will have positive entropies and conditional probabilities less than one (Richman et al., 2004). The algorithm to compute SampEn is in Table 4.2 (Alcaraz, Rieta, 2008).

The second entropy metric, termed Kullback Liebler Divergence (KLD), measures the relative entropy between two probability distributions, $p = \{p_k\}$ and $q = \{q_k\}$. In the present study, frequency distributions were analyzed instead of probability distributions, where k is a frequency and p_k is the normalized density at k, similarly for q_k (Quiroga et al., 2000). The KLD computed the relative entropy between a static frequency distribution constructed from the LL data, and a frequency distribution that changed with a sliding window over the TVIC. The hypothesis was that the frequency distribution from the HL and ML differed markedly from the LL distribution, thus KLD would peak during these tasks.

KLD is always positive and its magnitude is proportional to the difference between distributions; it is zero for identical distributions. KLD is computed by

$$KLD(p|q) = \sum_k p_k ln\left(\frac{p_k}{q_k}\right) \tag{4.5}$$

Results of the PDM on Various Metrics Several different metrics were presented, each with the ability to characterize the task-varying nature of the TVIC. The PDM was used to objectively evaluate which of the metrics most consistently peaked for tasks. This section presents those results, and identifies the metric which is best suited for online OFS change detection.

The standard deviation, skewness, and kurtosis metrics were all computed by a sliding window over each subject's TVIC. The span of the window was 18 seconds, providing an update of the metric each second. The length of the span was chosen for optimal peak production, as 18 seconds corresponds to the duration of the tasks; hence, at one point the span of the window would precisely overlap an entire task, resulting in the most extreme realization of the metric, i.e., the largest peak. The span may seem extensive, but it was necessary to smooth the metrics and resolve any task-induced trends.

For the following analyses, the δ and γ parameters of the PDM were optimized with respect to each subject and metric under evaluation. This strategy is consistent with operational requirements, as between-subjects variation is too profound to generalize parameter settings. An example of the standard deviation of the TVIC from E02 is plotted in Figure 4.1; the standard deviation is negated for interpretability and the peaks detected via the PDM are indicated by cross-hairs. All 8 tasks were detected.

Table 4.1 PDM algorithm

1. **Initialize:**
   - 1.1. Testing set, $\boldsymbol{S} = \{x_1, x_2, \ldots, x_n\}$
   - 1.2. δ, γ are user specified
   - 1.3. $L_{max} = \infty^-$, $I_{max} = 0$
   - 1.4. $L_{min} = \infty^+$, $I_{min} = 1$
   - 1.5. $B = 1$
   - 1.6. $\boldsymbol{P} = \emptyset$, $\boldsymbol{I} = \emptyset$
   - 1.7. $i = 1$
2. **Run Algorithm:**
   - 2.1. **while** $i < n$
   - 2.2. **if** $x_i > L_{max}$
   - 2.3. $L_{max} = x_i$, $I_{max} = i$
   - 2.4. **end if**
   - 2.5. **if** $x_i < L_{min}$
   - 2.6. $L_{min} = x_i$
   - 2.7. **end if**
   - 2.8. **if** $B == 1$
   - 2.9. **if** $x_i < L_{max} - \delta$ AND $L_{max} > \gamma$
   - 2.10. $\boldsymbol{P} \leftarrow \boldsymbol{P} \cup L_{max}$
   - 2.11. $\boldsymbol{I} \leftarrow \boldsymbol{I} \cup I_{max}$
   - 2.12. $B = 0$
   - 2.13. $L_{min} = x_i$
   - 2.14. **end if**
   - 2.15. **else if** $x_i > L_{min} + \delta$
   - 2.16. $L_{max} = x_i$
   - 2.17. $I_{max} = i$
   - 2.18. $B = 1$
   - 2.19. **end if/else**
   - 2.20. $i = i + 1$
   - 2.21. **end while**
3. **Output:** peaks and their indices are contained in sets $\boldsymbol{P}$ and $\boldsymbol{I}$, respectively

Table 4.2 SampEn algorithm

1. **Initialize:**
   - 1.1. N is the length of the record
   - 1.2. m and r are user defined parameters
   - 1.3. Form vectors of length m, $\boldsymbol{x}_m(1), \boldsymbol{x}_m(2), \ldots, \boldsymbol{x}_m(N-m+1)$
     - 1.3.1. $\boldsymbol{x}_m(i) = [x(i), x(i+1), \ldots, x(i+m-1)]\ \forall\ i$
   - 1.4. B_i, A_i, B^m, A^m, D are initialized to zero
   - 1.5. $i = 1, j = 1$
2. **Run Algorithm:**
   - 2.1. **while** $i \leq N - m + 1$
   - 2.2. $b = 0$
   - 2.3. **while** $j \leq N - m + 1$ AND $j \neq i$
   - 2.4. Select two vectors $\boldsymbol{x}_m(i)$ and $\boldsymbol{x}_m(j)$
   - 2.5. $$D = \max_{k=0,\ldots,m-1} |x(i+k) - x(j+k)|$$
   - 2.6. **if** $D < r$
   - 2.7. $b = b + 1$
   - 2.8. **end if**
   - 2.9. **end inner while**
   - 2.10. $$B_i = \frac{1}{N-m-1} b$$
   - 2.11. **end outer while**
   - 2.12. $$B^m = \frac{1}{N-m} \sum_{i=1}^{N-m} B_i$$
   - 2.13. Repeat steps 1-3, except for vectors of length $m + 1$. j ranges from 1 to $N - m$ and replace b with a, and B with A
3. **Output:**

$$SampEn = -ln\left[\frac{A^m}{B^m}\right]$$

SampEn was also computed by a sliding window, in this case, with a span of 10 seconds. The span was shorter because a second window was applied in addition to the first. SampEn exhibited brief, erratic outliers whose false peaks were not sufficiently smoothed by a single window. To compensate, a second window was applied over the first-windowed TVIC, and for each set of 10 points, it computed the area under the curve, i.e., approximated the integral. The double-windowing smoothed false peaks, as their short duration was not associated with large area under the curve. Figure 4.2 displays an example of this for A02. In (a), the first sliding window of SampEn was applied to the TVIC. Notice the presence of several short, narrow outlier peaks in contrast to the wider peaks exhibited during each task. Plot (b) is the integral from the second sliding window over the first windowed signal in (a); it demonstrates the dampening of outlier peaks and the correct detection of all 8 tasks with no false alarms. The parameters from SampEn were set as follows: r to .2 times the standard deviation of the record, m to 2, and the record length, N, to 2000 (Richman et al., 2004).

The last metric analyzed was the KLD which computed the relative entropy between two frequency distributions. The static frequency distribution was derived by averaging the frequencies of each 10 second epoch from the LL of an entire trial's TVIC; as before, the frequencies were computed via the DFT. The second distribution was computed in real-time from a sliding window over the subject's TVIC. The density of each frequency distribution was normalized, so the area under the distribution was one. After normalization, both the static baseline distribution and the real-time distribution were submitted to the KLD, for every second, to form a relative entropy signal.

The span of the sliding window for the KLD was again set to 10 seconds, because as with SampEn, the KLD necessitated a second window to dampen outlier peaks. As before, the second window slid over the initial windowed signal and computed the area under the curve for each set of 10 points. Figure 4.3 is an example for E02 where plot (a) is the KLD windowed over the TVIC and plot (b) is the integral from the second sliding window over the first windowed signal in (a). Similar to Figure 4.2, the double windowing smoothed outliers while amplifying the task-induced peaks to detect 7 of the 8 tasks.

The overall results of the PDM analysis are in Table 4.3. The proportion of correctly detected tasks is listed for each metric by trial. Average accuracy and false alarm rate is also reported in bold for each metric. According to the results, the standard deviation was the most proficient at producing peaks for each task. Using the standard deviation, the PDM detected 93.8% of the tasks with only .5 false alarms per trial on average. The standard deviation dominated all other metrics, i.e., for no subject or trial did another metric perform better. SampEn also had strong performance, detecting 81.3% of tasks on average, but it failed for subject F. The next most consistent metric was kurtosis, with 68.8% of the tasks correctly detected on average. Recall that kurtosis quantifies the shape of a distribution, meaning that with some success, task loads can be distinguished by the shape of the TVIC's amplitude distribution. KLD also utilized distributions, in this case, to quantify differences in entropy. The KLD metric detected 64.6% of tasks on average in the PDM.

Table 4.3 Results of the PDM on various metrics

	Std. Deviation	Skewness	Kurtosis	SampEn	KLD
A01	1.000	1.000	1.000	1.000	1.000
A02	1.000	0.250	0.750	1.000	0.750
E01	1.000	0.375	0.625	1.000	0.750
E02	1.000	0.750	0.625	1.000	0.875
F01	0.875	0.625	0.750	0.500	0.250
F02	0.750	0.625	0.375	0.375	0.250
Average	**0.938**	**0.604**	**0.688**	**0.813**	**0.646**
FA/Trial	**0.500**	**0.000**	**1.250**	**0.750**	**0.750**

Discussion of Metrics and the PDM The PDM was used as a tool to objectively evaluate how well each metric could detect change in task loads. The PDM analysis yielded that the windowed standard deviation of the TVIC was the best. By monitoring this signal and metric for each subject, a change detection scheme could accurately identify when the subject changes state.

The entropy metrics were also competitive in detecting tasks, but only for subjects A and E. Once again, subject F's psychophysiological signals proved difficult to accommodate, as they did not exhibit a recognizable entropy pattern. With data on more subjects, the entropy metrics may prove excellent in characterizing task-varying signals, on the other hand, it may be reaffirmed that entropies are largely subject-specific.

Yu first introduced the PDM's application to the UAV dataset. He presented task detection results from the PDM on the windowed standard deviation of a subject's F_z signal (2009). The span of the sliding window was set to 18 seconds and the PDM parameters, δ and γ, were fixed. In order to make a fair comparison with these findings, the PDM analysis was repeated, this time with fixed parameters on the standard deviation of a subject's TVIC. The outcome of the comparison is in Table 4.4.

The average accuracy using a subject's TVIC was 2% higher than using their F_z signal. While this is not a significant improvement, the PDM on the TVIC yielded less than half the false alarms on average when compared to the F_z signal.

The results support the notion of using independent components instead of the raw psychophysiological signals for OFS change detection. Independent components are less complex and less noisy, and thus result in fewer false alarms. Not only do these results confirm ICA's ability to extract artifacts, they directly demonstrate ICA's ability to facilitate the detection of change in a subject's cognitive state.

The PDM is a valuable tool, but its role should not be extrapolated to that of an online change detector. In the context of the present study, a peak is only defined once a metric reverses direction, and begins descending back to the LL baseline. In other words, a peak is only detected once the cognitive load returns to normal, long after a task has ended. Therefore, the PDM cannot be a true online change detector, as it can only detect tasks in retrospect.

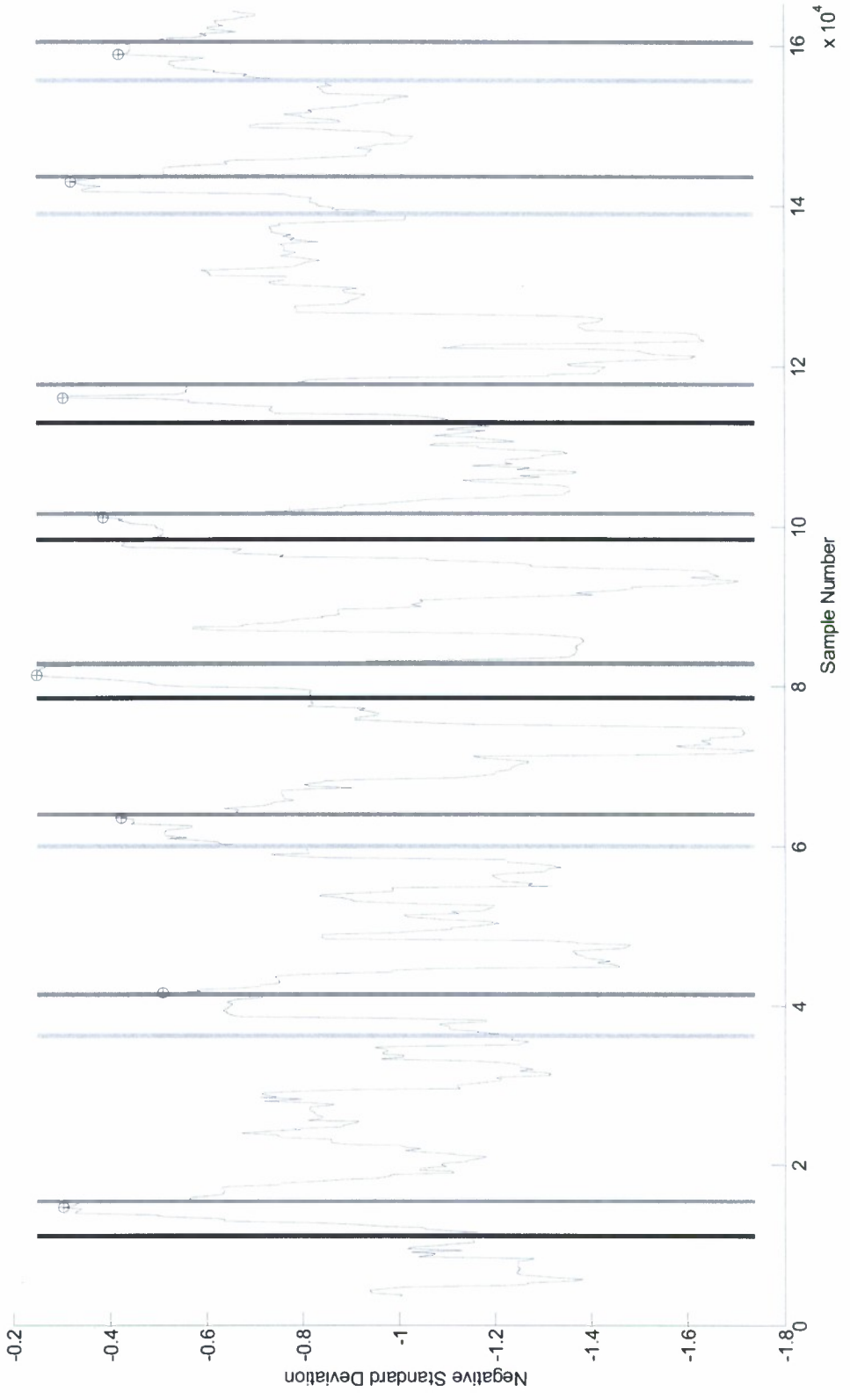

Figure 4.1 PDM on the standard deviation of the TVIC of E02

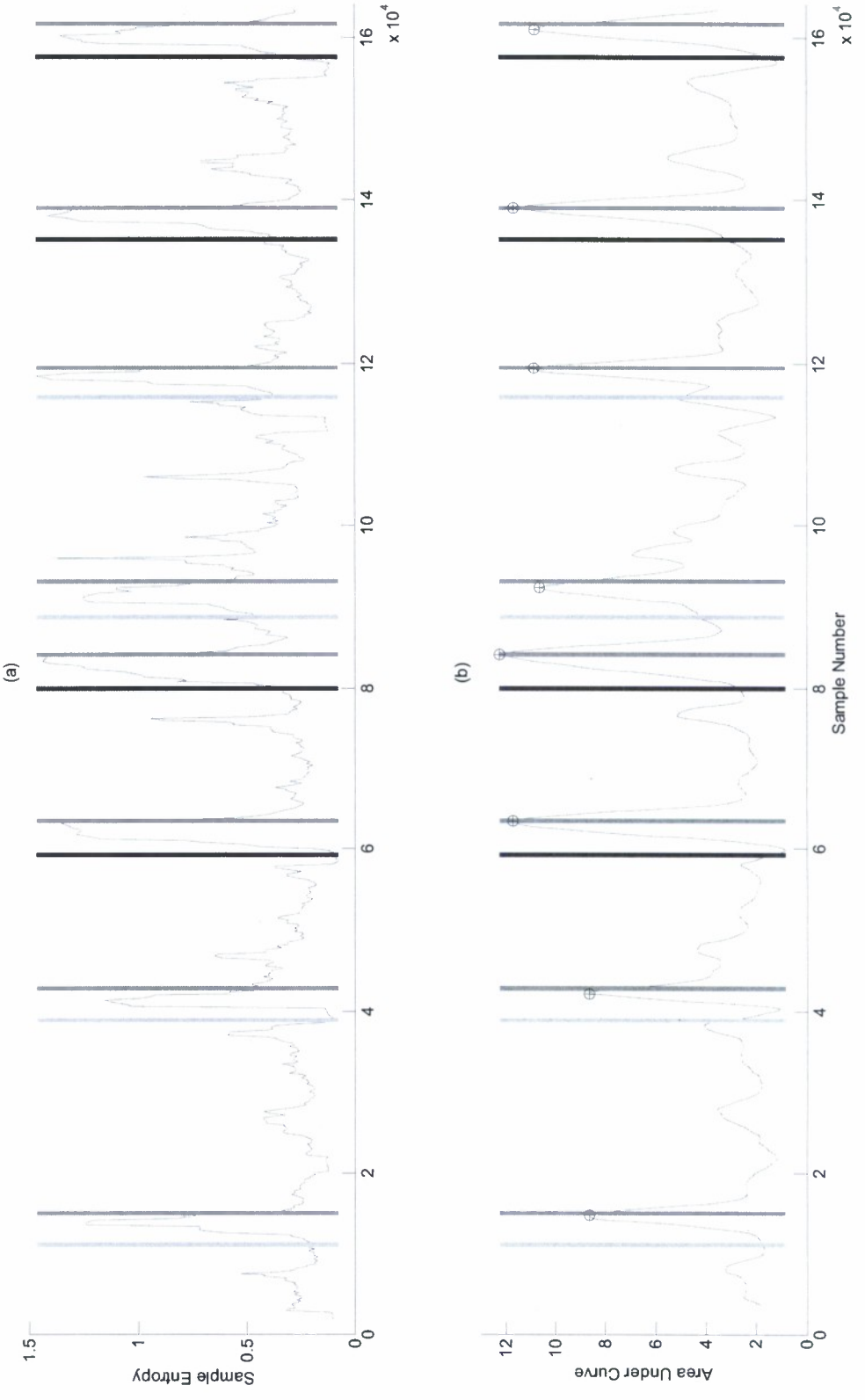

Figure 4.2 PDM on the SampEn of the TVIC of A02 (a) first sliding window (b) second window of integral on (a)

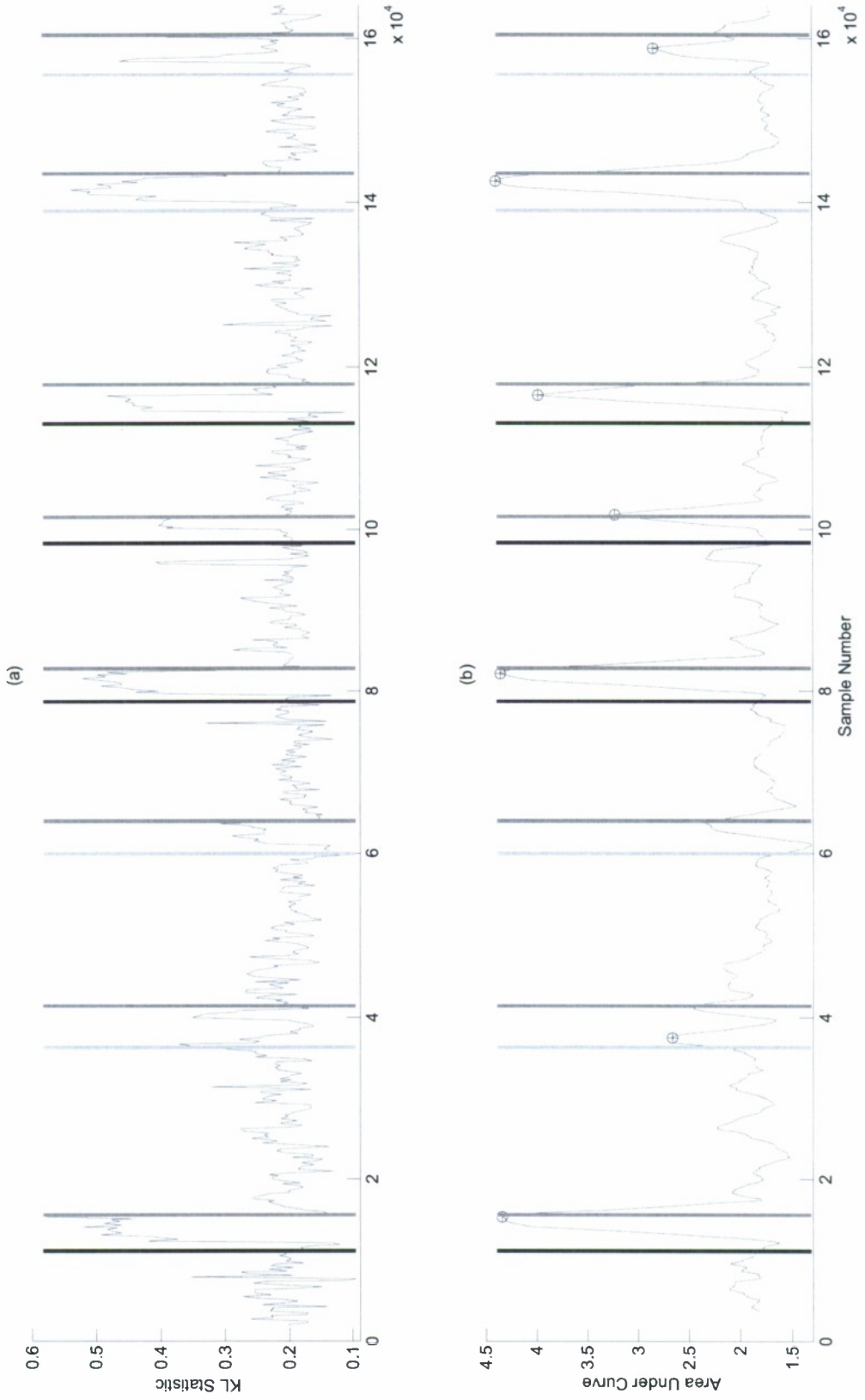

Figure 4.3 PDM on the KLD of the TVIC of E02 (a) first sliding window (b) second window of integral on (a)

Table 4.4 Results of the PDM on the standard deviation of F_z and TVIC

	F_z	TVIC
A01	1.000	1.000
A02	0.875	0.750
E01	1.000	1.000
E02	0.750	0.875
F01	1.000	0.875
F02	0.500	0.750
Average	**0.854**	**0.875**
FA/Trial	**2.000**	**0.833**

The PDM does provide an interesting model, one that could be modified into a real-time scheme that detects a metric's task-induced *trends*. After identifying the standard deviation as the optimal metric to characterize a subject's TVIC, the next step is to employ the metric in a real-time OFS change detector. The following section introduces a novel algorithm, loosely based on the PDM, which monitors and detects task-induced trends instead of peaks.

The Trend Detection Method A true online classifier cannot benefit from hindsight, meaning the information used to classify an epoch of data must come from that epoch and/or the information that preceded it. The PDM does not qualify as a true online change detector because it requires information about what transpires *after* an epoch in order to retroactively detect peaks. An online classifier must also be instantaneous, and most importantly, it must be accurate.

The *trend detection method* (TDM) introduced in this section was developed to meet these requirements and to address the deficiencies of the PDM. The TDM assumes that the ML and HL will cause a windowed metric to monotonically increase (or decrease) from the LL baseline. This behavior was observed when monitoring the standard deviation of a subject's TVIC. The TDM identifies and detects these task-induced trends in order to detect real-time changes in OFS.

The keys to trend detection are to identify significant trends through noise and to do so for non-stationary signals, where trends can begin at different magnitudes over time. The TDM addresses these challenges via an algorithm, presented in Table 4., which has two main components: an adaptive threshold and a trend detector. The adaptive threshold adjusts to non-stationary behavior and the second component, the trend detector, identifies task-induced trends. Detecting trends accomplishes two things: first, coupled with information on the signal's magnitude, a trend helps discriminate between true task-induced increases and irrelevant transient spikes. Second, depending on the length of the trend, tasks occurring below the threshold can be identified and adapt the threshold accordingly.

The TDM judges a signal, x_i, at time i against two criteria: its magnitude relative to the adaptive threshold and how well the previous points, $x_{i-1}, x_{i-2}, \ldots, x_{i-n}$, have

trended. Trends are identified by counting the number of positive "slopes" between consecutive points over a period of time, where τ is the count. A positive slope is defined when $x_i > x_{i-1}$. The slopes need not be consecutive since noise may cause temporary reversals in the trend; reversals will only dissolve a trend if, after a period of time, no point falls above the trend's last point, say, x_i. This time period is defined by the parameter R_{lim}. If, however, R_{lim} is not exceeded and some point, $x_{i+n} > x_i$, then $\tau = \tau + 1$ and the trend detection continues. When $\geq T_{low}$, the second phase of the algorithm is commenced.

Once an emerging trend is identified, the last value in the trend, x_i, is compared to the threshold, Z, to determine if a task-induced cognitive load has occurred. If $x_i > Z$, the trend is classified as task-induced. In contrast, if $x_i < Z$, the emerging trend remains unclassified but continues to be monitored until $\tau > T_{up}$, where T_{up} is a newly defined limit. In summary, the second phase of the algorithm classifies trends in one of two ways: the first is when the trend outright breaches Z, and the second, is when the trend is occurring below Z, but is sustained long enough to be deemed task-induced. T_{up} quantifies "long enough" by accounting for the distance x_i falls below Z; the further below Z, the longer trending must continue in order to be classified. The motivation here is to account for non-stationarity while preventing the classification of trends not associated with a task load.

If a task-induced trend is classified below the current Z, the threshold updates by an EWMA that weights the magnitudes of previously classified trends with the most recent trend, through a smoothing parameter λ. This mechanism adapts the threshold so it can detect task-induced trends occurring at different magnitudes in the future. These mechanisms of the TDM are all detailed in Table 4.5.

Results of the TDM on the Standard Deviation of TVIC The TDM was conducted on the standard deviation of an 18 second sliding window over each subject's TVIC; this is the same setup as the PDM analysis. Normally, the parameters of the TDM are calibrated during training. However, for the purpose of comparing results to the PDM, the parameters were calibrated on the testing data. In general, the parameters were set within the following ranges: T_{low} [4,6], L [1.2,1.4], β [8,9]. The α and R_{lim} parameters were both held constant at .3 and 5, respectively. Figure 4.4 is a plot of the TDM of E01; 8 task-induced trends were detected, with one false alarm. Notice the instances a trend was identified more than once for the same task; this only indicates the TDM detected the task early and continued to detect further trending as the task proceeded.

The results of the TDM analysis for each trial are in Table 4.5. The accuracy is defined as the proportion of tasks that were correctly detected. The average accuracy of the TDM was 85.4%, with over one false alarm per trial on average. Lastly, the tasks were detected, on average, 13 seconds after their onset, a comparatively quick detection when compared to the half minute delays in control charting.

Table 4.5 TDM algorithm

1. **Initialize:**
   - 1.1. Testing set, $\mathbf{S} = \{x_1, x_2, \ldots, x_n\}$
   - 1.2. T_{low}, R_{lim}, λ, β, and L are all defined during training
   - 1.3. $\tau = 0$, $v = 0$
   - 1.4. $T_{up} = \infty^+$
   - 1.5. $L_{max} = \mu_X$, where μ_X is the mean of **S** or a training set X
   - 1.6. $Z_o = \mu_X + L\sigma_X$, where σ_X is the standard deviation **S** or of a training set X
   - 1.7. $\mathbf{V} = \emptyset$, $\mathbf{I} = \emptyset$
   - 1.8. $i = 1$
2. **Run Algorithm:**
   - 2.1. **do**
   - 2.2. $i = i+1$
   - 2.3. **if** $x_i > x_{i-1}$
   - 2.4. **if** $x_i > L_{max}$
   - 2.5. $L_{max} = x_i$
   - 2.6. $\tau = \tau + 1$
   - 2.7. $v = 0$
   - 2.8. **else** $v = v + 1$ **end if/else**
   - 2.9. **else** $v = v + 1$ **end if/else**
   - 2.1. **while** ($v < R_{lim}$ AND $\tau < T_{low}$ AND $i < t$)
   - 2.2. **if** $v \geq R_{lim}$
   - 2.3. $\tau = 0$
   - 2.4. $L_{max} = x_i$
   - 2.5. **else if** $x_i \geq Z$ OR $\tau \geq T_{up}$
   - 2.6. $\mathbf{V} \leftarrow \mathbf{V} \cup x_i$
   - 2.7. $\mathbf{I} \leftarrow \mathbf{I} \cup i$
   - 2.8. $\tau = 0$, $v = 0$
   - 2.9. $T_{up} = \infty^+$
   - 2.10. $L_{max} = \mu_X$
   - 2.11. $Z = min\{ Z_o, Z(1-\lambda) + \lambda L_{max} \}$

Table 4.5 Results of the TDM analysis

Trial	Proportion Detected
A01	1.00
A02	.750
E01	1.00
E02	.875
F01	.750
F02	.750
Average	**.854**
FA/Trial	**1.166**

Discussion of the TDM Overall, the results indicate that the TDM detected changes in cognitive load accurately, quickly, and most importantly, in real-time. Pattern recognition methods that classify epochs into discrete categories of cognition are not conducive to real-time change detection; it is unclear when a subject changes state. In contrast, the TDM, in combination with a subject's TVIC, can clearly indicate when a change has occurred.

The principle shortcomings of the TDM are its inability to signal when a task has ended and the number of parameters that must be calibrated. In adaptive aiding, it is necessary for the TDM to not only identify the beginning of a task, but also indicate the task's conclusion so aiding can be withdrawn. This capability can be added to the algorithm by monitoring trends that reverse to the baseline once a task completes. The second drawback is the number of parameters that need to be set for the TDM, five in all. In the present analysis, only three parameters required adjustment across trials, but even setting these parameters is not straightforward. This process could be formalized by conducting a design of experiments on training data, scientifically determining the optimal parameter settings for each subject.

These suggestions to remedy the shortcomings of the TDM were not pursued further, since the scope of this study is a proof-of-concept rather than an optimization of any single method. The goal of employing adaptive aiding from psychophysiological signals will require more than a single method. The TDM and the TVIC have been developed to augment the arsenal of schemes which identify changing cognitive load, and it is hoped that in concert with these established methods, future adaptive aiding schemes can be realized.

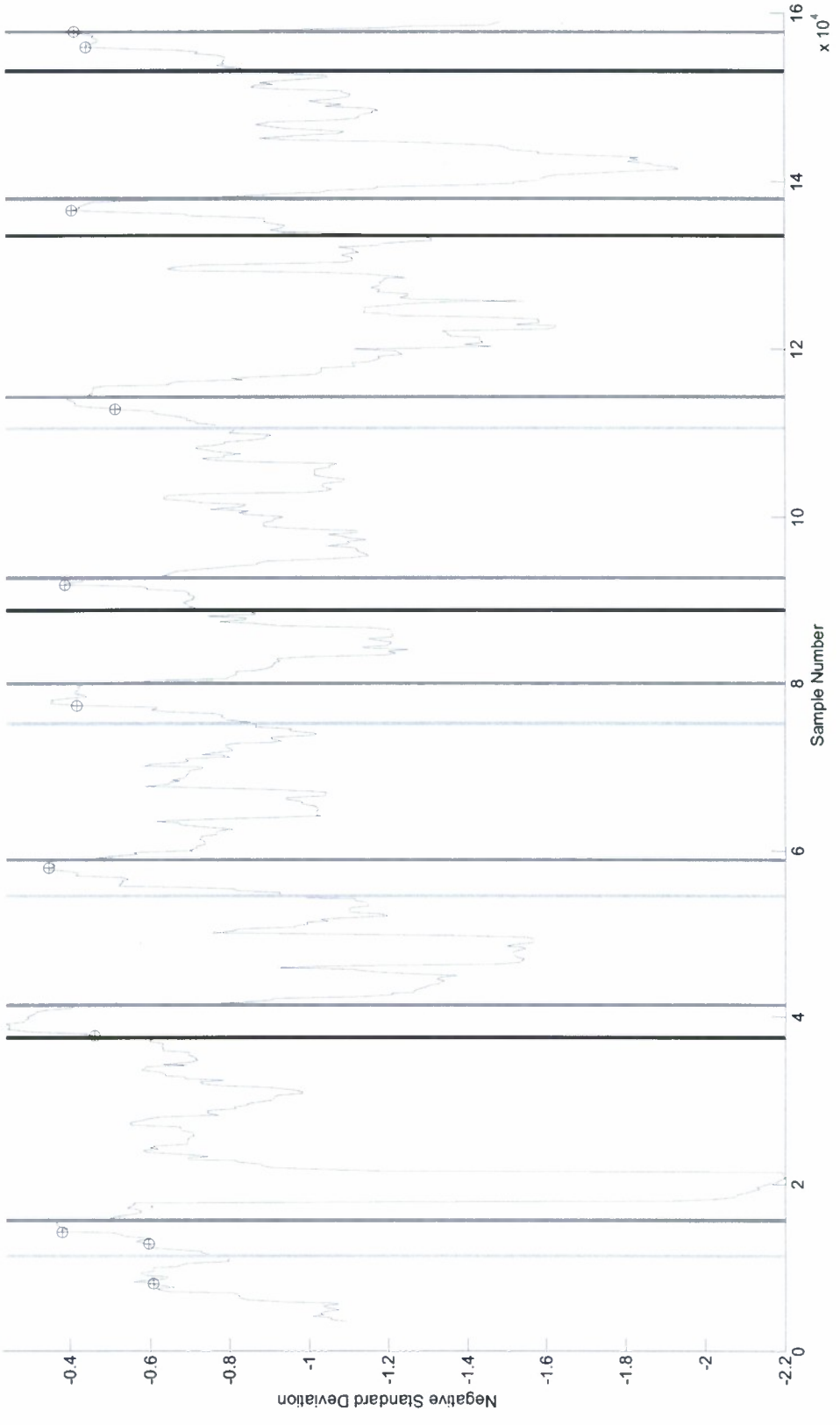

Figure 4.4 TDM on the standard deviation of the TVIC of E01

APPENDIX A
TASK LOAD MEANS ANALYSIS

Table A.1 Means of features across task loads for A01

	Theta	Alpha	Beta
LL	1.182	0.637	0.988
ML	1.246	0.609	0.809
HL	1.213	0.572	0.614

Table A.2 Means of features across task loads for E01

	Theta	Alpha	Beta
LL	1.685	0.569	1.023
ML	1.564	0.447	0.654
HL	1.493	0.635	0.870

Table A.3 Means of features across task loads for F01

	Theta	Alpha	Beta
LL	0.987	0.332	0.585
ML	0.925	0.277	0.414
HL	1.026	0.263	0.439

Table A.4 Means of features for A01, E01, and F01 averaged across task loads

	Theta	**Alpha**	**Beta**
LL	1.283	0.514	0.866
ML	1.237	0.435	0.614
HL	1.241	0.487	0.639

APPENDIX B
FALSE ALARM DATA FROM CONTROL CHARTS

Table B.1 False alarms of EWMA-Shewart control charts

	VEOG Theta	F_z Theta	F_z Alpha	F_7 Theta	O_2 Alpha	P_z Alpha	Average
A01	1	2	1	2	0	1	**1.167**
E01	0	0	1	1	0	0	**0.333**
F01	1	0	1	0	1	0	**0.500**
Average	**0.667**	**0.667**	**1.000**	**1.000**	**0.333**	**0.333**	

Table B.2 False alarms of time series residuals control charts

	HEOG Beta	F_z Beta	P_z Beta	T_5 Beta	Average
A01	1	1	2	1	**1.250**
E01	2	1	1	1	**1.250**
F01	4	4	3	1	**3.000**
Average	**2.333**	**2.000**	**2.000**	**1.000**	

Table B.3 False alarms of MCEMWA control charts

	F_z Beta	P_z Beta	T_5 Beta	Average
A01	3	3	3	**3.000**
E01	2	2	1	**1.667**
F01	2	4	2	**2.667**
Average	**2.333**	**3.000**	**2.000**	

Table B.4 False alarms of various multivariate control charts

	Hotelling-T^2 (Select)	MEWMA (Select)	Hotelling-T^2 (PCA)	MEWMA (PCA)	Average
A01	4	2	4	3	**3.250**
E01	4	1	2	2	**2.250**
F01	6	4	4	2	**4.000**
Average	**4.667**	**2.333**	**3.333**	**2.333**	

APPENDIX C
PRINCPLE COMPONENT FACTOR PATTERN ANALYSIS

Table C.1 Principle component factor pattern of A01

		P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
VEOG	delta	0.301	-0.696	0.295	-0.036	0.053	-0.460	0.098	-0.016	0.017	-0.026
	theta	0.353	-0.634	0.308	-0.039	0.029	-0.463	0.236	0.037	0.109	-0.095
	alpha	0.466	-0.583	0.024	-0.029	-0.039	-0.368	0.359	0.003	0.074	-0.045
	beta	0.429	-0.093	-0.667	-0.106	0.136	-0.049	0.454	-0.017	-0.197	0.045
	gamma	0.338	0.168	-0.680	-0.083	0.118	-0.002	0.426	-0.034	-0.237	0.101
HEOG	delta	0.030	-0.786	0.129	-0.129	-0.074	-0.082	-0.314	-0.024	0.044	0.042
	theta	0.159	-0.826	-0.147	-0.162	0.019	0.041	-0.296	-0.007	0.080	-0.017
	alpha	0.191	-0.695	-0.512	-0.191	0.171	0.202	-0.161	0.039	0.067	-0.066
	beta	0.116	-0.645	-0.581	-0.195	0.271	0.175	-0.116	0.069	0.075	-0.081
	gamma	0.148	-0.632	-0.573	-0.167	0.290	0.157	-0.105	0.076	0.071	-0.071
F_z	delta	0.355	-0.207	-0.005	0.497	0.140	0.086	-0.157	-0.027	0.036	0.604
	theta	0.230	-0.235	-0.149	0.014	-0.491	0.021	-0.288	-0.324	-0.568	-0.028
	alpha	0.106	-0.237	0.250	-0.191	-0.120	0.437	0.300	-0.560	0.235	0.080
	beta	0.804	0.275	0.027	-0.103	0.036	0.071	-0.025	-0.018	0.063	0.030
	gamma	0.808	0.432	-0.031	-0.115	0.034	-0.039	-0.157	-0.003	0.058	-0.006
F_7	delta	0.429	-0.471	0.114	0.324	0.138	-0.236	-0.020	-0.089	-0.063	0.360
	theta	0.398	-0.637	0.009	-0.013	-0.198	-0.157	-0.168	-0.161	-0.153	-0.049
	alpha	0.450	-0.356	0.070	-0.017	-0.312	0.256	0.198	-0.305	0.303	-0.017
	beta	0.745	0.306	-0.172	-0.017	-0.013	0.095	0.022	-0.024	0.052	0.094
	gamma	0.776	0.337	-0.096	-0.040	0.010	-0.017	-0.084	-0.019	0.025	0.135

Table C.1—Continued

P_z	**delta**	0.378	-0.155	0.017	0.759	0.167	0.182	-0.027	-0.048	-0.043	-0.104
	theta	0.287	-0.341	0.042	-0.022	-0.632	0.140	0.177	0.180	-0.235	-0.013
	alpha	0.174	-0.250	0.590	-0.313	0.320	0.347	0.115	0.042	-0.257	-0.035
	beta	0.852	0.090	0.213	-0.110	0.084	-0.004	-0.034	0.062	-0.025	-0.072
	gamma	0.790	0.421	-0.036	-0.149	-0.041	-0.064	-0.194	-0.030	0.070	-0.065
T_5	**delta**	0.338	-0.120	0.002	0.760	0.183	0.089	0.038	-0.013	-0.073	-0.238
	theta	0.328	-0.288	-0.074	0.223	-0.470	0.160	-0.024	0.178	0.171	-0.181
	alpha	0.382	-0.146	0.504	-0.350	0.234	0.298	0.057	0.020	-0.152	-0.019
	beta	0.845	0.329	0.020	-0.047	0.011	-0.079	-0.010	0.067	0.019	-0.095
	gamma	0.768	0.509	-0.116	-0.125	-0.053	-0.063	-0.061	-0.015	0.003	-0.068
O_2	**delta**	0.301	-0.087	0.114	0.761	0.222	0.151	0.066	-0.055	-0.038	-0.256
	theta	0.216	-0.197	-0.090	0.195	-0.517	0.251	0.143	0.504	0.125	0.179
	alpha	0.291	-0.370	0.517	-0.198	0.214	0.262	0.056	0.335	-0.155	0.088
	beta	0.852	0.021	0.213	-0.040	0.038	-0.041	-0.024	0.080	0.007	0.001
	gamma	0.827	0.395	-0.011	-0.118	-0.028	-0.053	-0.153	-0.020	0.067	-0.049

Table C.2 Principle component factor pattern of E01

		P1	P2	P3	P4	P5	P6
VEOG	delta	-0.5189	-0.6534	0.144	-0.2419	0.0522	0.2724
	theta	-0.4243	-0.7641	0.0467	-0.159	0.1439	0.2716
	alpha	-0.433	-0.6961	0.1683	-0.0844	0.0845	0.3606
	beta	-0.6895	-0.3628	0.0832	-0.3572	0.1927	0.0261
	gamma	-0.7851	0.0523	0.2799	-0.2339	0.272	-0.0508
HEOG	delta	-0.336	-0.4305	0.355	0.5134	0.0103	0.2619
	theta	-0.3309	-0.384	0.4095	0.6438	-0.0397	0.0569
	alpha	-0.3289	-0.2771	0.4425	0.6549	0.0255	-0.2625
	beta	-0.5119	-0.1063	0.4952	0.3784	0.2077	-0.3599
	gamma	-0.6706	0.1094	0.4881	0.0588	0.2694	-0.2786
F_z	delta	-0.6035	-0.4614	-0.1269	-0.2895	-0.1911	-0.2703
	theta	-0.3918	-0.5648	-0.4679	-0.114	-0.076	-0.3507
	alpha	-0.4978	-0.1078	-0.5945	0.3151	0.2686	-0.0735
	beta	-0.8873	0.2272	-0.1378	0.0452	0.1014	0.0789
	gamma	-0.9107	0.2738	0.1111	-0.1049	0.0996	0.0166
F_7	delta	-0.7248	-0.4215	0.0103	-0.2809	0.0081	-0.1126
	theta	-0.4696	-0.6763	-0.3183	-0.1975	0.0708	-0.2488
	alpha	-0.7304	-0.2067	-0.4707	0.1129	0.1365	0.0058
	beta	-0.9062	0.2418	-0.0412	0.0034	0.0692	0.104
	gamma	-0.8862	0.2518	0.1427	-0.0794	0.1388	0.0524

Table C.2—Continued

P_z	**delta**	-0.6402	-0.08	0.1507	-0.1932	-0.5275	-0.0622
	theta	-0.4799	-0.2524	-0.2622	0.3264	-0.4346	-0.0246
	alpha	-0.4693	-0.0026	-0.5046	0.4817	0.1778	0.0789
	beta	-0.9078	0.2354	-0.1689	0.0688	0.0452	0.0739
	gamma	-0.911	0.3084	0.0667	-0.1007	0.0427	0.0329
T_5	**delta**	-0.6846	0.1718	0.2425	-0.1408	-0.3044	-0.051
	theta	-0.7532	0.1347	-0.1896	0.0684	-0.1755	-0.1948
	alpha	-0.7807	0.2635	-0.3097	0.2122	0.0648	0.0046
	beta	-0.9077	0.3354	0.0736	-0.097	0.0588	0.045
	gamma	-0.8565	0.3032	0.189	-0.1725	0.0366	-0.0144
O_2	**delta**	-0.6391	-0.0012	0.2664	-0.0898	-0.5272	0.0437
	theta	-0.5422	0.0021	-0.1065	0.3426	-0.4705	0.1319
	alpha	-0.6947	0.2539	-0.3355	0.2498	-0.0538	0.1979
	beta	-0.8691	0.267	-0.0338	-0.0628	-0.0304	0.0687
	gamma	-0.8201	0.2839	0.0569	-0.1676	-0.008	0.0435

Table C.3 Principle component factor pattern of F01

		P1	P2	P3	P4	P5	P6
VEOG	delta	0.6481	0.4212	-0.3956	0.2711	-0.2277	0.1072
	theta	0.6769	0.4747	-0.4191	0.156	-0.093	0.166
	alpha	0.6629	0.4442	-0.4431	0.1988	-0.1474	0.1474
	beta	0.7153	0.4356	-0.3604	0.0943	-0.1592	-0.0201
	gamma	0.7047	0.3298	-0.2268	-0.0884	-0.1227	-0.2113
HEOG	delta	0.4053	0.4785	-0.007	-0.2073	0.5887	0.0549
	theta	0.3171	0.4229	-0.0119	-0.2102	0.7706	0.1341
	alpha	0.327	0.401	0.0166	-0.2174	0.7216	0.0349
	beta	0.7429	0.0939	-0.0473	-0.2071	0.2727	-0.3319
	gamma	0.8429	-0.1172	-0.0888	-0.1152	0.1073	-0.3839
F_z	delta	0.6776	0.2758	0.1412	-0.0529	-0.2137	0.1642
	theta	0.7061	0.1645	0.0022	0.0995	-0.1612	0.3521
	alpha	0.5967	0.0709	0.2467	0.6371	-0.0346	-0.0227
	beta	0.9084	-0.1756	-0.129	-0.0609	-0.1039	-0.1901
	gamma	0.8487	-0.3597	-0.1385	-0.1057	-0.0291	-0.2219
F_7	delta	0.792	0.0362	0.0809	-0.0665	-0.0444	-0.198
	theta	0.8477	0.1565	-0.18	0.1385	-0.0355	0.0885
	alpha	0.7383	0.1694	-0.0532	0.4833	0.0014	-0.0765
	beta	0.8639	-0.11	-0.2735	-0.0184	-0.0179	-0.0425
	gamma	0.8661	-0.2598	-0.2619	-0.0743	-0.0073	-0.078

Table C.3—Continued

P_z	**delta**	0.5817	0.3189	0.4502	-0.3692	-0.3161	-0.0487
	theta	0.5812	-0.005	0.3404	-0.0409	0.0119	0.2805
	alpha	0.3621	-0.1135	0.4305	0.6667	0.2255	-0.1346
	beta	0.8372	-0.4329	-0.0042	-0.0682	0.1052	-0.1091
	gamma	0.7565	-0.5424	-0.0852	-0.0934	0.0703	-0.0969
T_5	**delta**	0.5121	0.3098	0.5481	-0.4078	-0.2885	-0.1013
	theta	0.6887	-0.0558	0.4622	-0.1293	-0.0936	0.0457
	alpha	0.4858	-0.068	0.5637	0.483	0.1444	-0.1009
	beta	0.6636	-0.559	-0.0096	-0.1271	0.1269	0.1912
	gamma	0.5974	-0.6202	-0.1568	-0.1055	0.1261	0.2266
O_2	**delta**	0.4669	0.3024	0.5679	-0.383	-0.2818	-0.1089
	theta	0.3816	0.222	0.4075	-0.2627	-0.013	0.3271
	alpha	0.2958	-0.0041	0.5662	0.5619	0.1806	0.0296
	beta	0.7242	-0.4314	0.0924	-0.0718	0.0529	0.3149
	gamma	0.6222	-0.5409	-0.0505	-0.0899	-0.0604	0.3376

REFERENCES

Alcaraz, R., & Rieta, J. J. (2008). Wavelet bidomain sample entropy analysis to predict spontaneous termination of atrial fibrillation. *Physiological Measurement, 29*(1), 65-80.

Brookings, J. B., Wilson, G. F., & Swain, C. R. (1996). Psychophysiological responses to changes in workload during simulated air traffic control. *Biological Psychology, 42*(3), 361-377.

Bruzzo, A. A., Gesierich, B., Santi, M., Tassinari, C. A., Birbaumer, N., & Rubboli, G. (2008). Permutation entropy to detect vigilance changes and preictal states from scalp EEG in epileptic patients. A preliminary study. *Neurological Sciences, 29*(1), 3-9.

Byrne, E. A., & Parasuraman, R. (1996). Psychophysiology and adaptive automation. *Biological Psychology, 42*(3), 249-268.

Fairclough, S. H., Venables, L., & Tattersall, A. (2005). The influence of task demand and learning on the psychophysiological response. *International Journal of Psychophysiology, 56*(2), 171-184.

Fournier, L. R., Wilson, G. F., & Swain, C. R. (1999). Electrophysiological, behavioral, and subjective indexes of workload when performing multiple tasks: manipulations of task difficulty and training. *International Journal of Psychophysiology, 31*(2), 129-145.

Freeman, F. G., Mikulka, P. J., Prinzel, L. J., & Scerbo, M. W. (1999). Evaluation of an adaptive automation system using three EEG indices with a visual tracking task. *Biological Psychology, 50*(1), 61-76.

Gevins, A., & Smith, M. (2003). Neurophysiological measures of cognitive workload during human-computer interaction. *Theoretical Issues in Ergonomics Science, 4*, 18.

Gevins, A., Smith, M. E., Leong, H., McEvoy, L., Whitfield, S., Du, R., et al. (1998). Monitoring working memory load during computer-based tasks with EEG pattern recognition methods. *Human Factors, 40*(1), 79-91.

Gevins, A., Smith, M. E., McEvoy, L., & Yu, D. (1997). High-resolution EEG mapping of cortical activation related to working memory: Effects of task difficulty, type of processing, and practice. *Cerebral Cortex, 7*(4), 374-385.

Hankins, T. C., & Wilson, G. F. (1998). A comparison of heart rate, eye activity, EEG and subjective measures of pilot mental workload during flight. *Aviation Space and Environmental Medicine, 69*(4), 360-367.

He, P., Wilson, G., Russell, C., & Gerschutz, M. (2007). Removal of ocular artifacts from the EEG: a comparison between time-domain regression method and adaptive filtering method using simulated data. *Medical & Biological Engineering & Computing, 45*(5), 495-503.

Hyvarinen, A., & Oja, E. (2000). Independent component analysis: algorithms and applications. *Neural Networks, 13*(4-5), 411-430.

Inouye, T., Shinosaki, K., Iyama, A., Matsumoto, Y., Toi, S., & Ishihara, T. (1994). POTENTIAL FLOW OF FRONTAL MIDLINE THETA-ACTIVITY DURING A MENTAL TASK IN THE HUMAN ELECTROENCEPHALOGRAM. *Neuroscience Letters, 169*(I-2), 145-148.

Ishii, R., Shinosaki, K., Ukai, S., Inouye, T., Ishihara, T., Yoshimine, T., et al. (1999). Medial prefrontal cortex generates frontal midline theta rhythm. *Neuroreport, 10*(4), 675-679.

Jackson, J. E. (1980). PRINCIPAL COMPONENTS AND FACTOR-ANALYSIS .1. PRINCIPAL COMPONENTS. *Journal of Quality Technology, 12*(4), 201-213.

Jung, T. P., Makeig, S., Stensmo, M., & Sejnowski, T. J. (1997). Estimating alertness from the EEG power spectrum. *Ieee Transactions on Biomedical Engineering, 44*(1), 60-69.

Klimesch, W., Schmike, H., & Pfurtscheller, G. (1993). Alpha frequency, cognitive load, and memory performance. *Brain Topography, 5*, 10.

Lowry, C. A., & Montgomery, D. C. (1995). A Review of Multivariate Control Charts. *IIE Transactions, 27*, 10.

Lowry, C. A., Woodall, W. H., Champ, C. W., & Rigdon, S. E. (1992). A MULTIVARIATE EXPONENTIALLY WEIGHTED MOVING AVERAGE CONTROL CHART. *Technometrics, 34*(1), 46-53.

Makeig, S., Bell, A., Jung, T.-P., & Sejnowski, T. (1996). Independent Component Analysis of Electroencaphalographic Data. *Advances in Neural Information Processing Systems 8*, 6.

Makeig, S., & Inlow, M. (1993). LAPSES IN ALERTNESS - COHERENCE OF FLUCTUATIONS IN PERFORMANCE AND EEG SPECTRUM. *Electroencephalography and Clinical Neurophysiology, 86*(1), 23-35.

Mastrangelo, C. M., Runger, G. C., & Montgomery, D. C. (1996). Statistical process monitoring with principal components. *Quality and Reliability Engineering International, 12*(3), 203-210.

Montgomery, D. (2009). *Statistical Quality Control* (6th edition ed.): John Wiley & Sons, Inc.

Montgomery, D. C., Jennings, C. L., & Kulachi, M. (2008). *Introduction to Time Series Analysis and Forecasting*. Hoboken: John Wiley & Sons.

Montgomery, D. C., & Mastrangelo, C. M. (1991). SOME STATISTICAL PROCESS-CONTROL METHODS FOR AUTOCORRELATED DATA. *Journal of Quality Technology, 23*(3), 179-193.

Phillips, C., Parr, J., & Riskin, E. (2008). *Signals, Systems, and Transforms* (4 ed.). Upper Saddle River: Prentice Hall.

Pope, A. T., Bogart, E. H., & Bartolome, D. S. (1995). BIOCYBERNETIC SYSTEM EVALUATES INDEXES OF OPERATOR ENGAGEMENT IN AUTOMATED TASK. *Biological Psychology, 40*(1-2), 187-195.

Prinzel, L. J., Freeman, F. C., Scerbo, M. W., Mikulka, P. J., & Pope, A. T. (2000). A closed-loop system for examining psychophysiological measures for adaptive task allocation. *International Journal of Aviation Psychology, 10*(4), 393-410.

Quiroga, R. Q., Arnhold, J., Lehnertz, K., & Grassberger, P. (2000). Kulback-Leibler and renormalized entropies: Applications to electroencephalograms of epilepsy patients. *Physical Review E, 62*(6), 8380-8386.

Richman, J. S., Lake, D. E., & Moorman, J. R. (2004). Sample entropy *Numerical Computer Methods, Pt E* (Vol. 384, pp. 172-184).

Richman, J. S., & Moorman, J. R. (2000). Physiological time-series analysis using approximate entropy and sample entropy. *American Journal of Physiology-Heart and Circulatory Physiology, 278*(6), H2039-H2049.

Scranton, R., Runger, G. C., Montgomery, D. C., & Keats, B. J. (1996). Efficient Shift Detection Using Multivariate Exponentially-Weighted Moving Average Control Charts and Principle Components. *Quality and Reliability Engineering International, 12*, 6.

Smith, M. E., Gevins, A., Brown, H., Karnik, A., & Du, R. (2001). Monitoring task loading with multivariate EEG measures during complex forms of human-computer interaction. *Human Factors, 43*(3), 366-380.

Stone, J. V. (2002). Independent component analysis: an introduction. [Review]. *Trends in Cognitive Sciences, 6*(2), 59-64.

Tou, J., & Gonzalez, R. (1974). *Pattern Recognition Principles*. Reading, MA: Addison-Wesley.

Vigario, R., Sarela, J., Jousmaki, V., Hamalainen, M., & Oja, E. (2000). Independent component approach to the analysis of EEG and MEG recordings. *Ieee Transactions on Biomedical Engineering, 47*(5), 589-593.

Wilson, G. F., & Fisher, F. (1991). THE USE OF CARDIAC AND EYE BLINK MEASURES TO DETERMINE FLIGHT SEGMENT IN F4-CREWS. *Aviation Space and Environmental Medicine, 62*(10), 959-962.

Wilson, G. F., & Fisher, F. (1995). COGNITIVE TASK CLASSIFICATION BASED UPON TOPOGRAPHIC EEG DATA. [Article]. *Biological Psychology, 40*(1-2), 239-250.

Wilson, G. F., & Russell, C. (1999). *Operator Functional State Classification Using Neural Networks with Combined Physiological and Performance Features*. Paper presented at the Proceedings of the Human Factors and Ergonomics Society 43rd Annual Meeting.

Wilson, G. F., & Russell, C. A. (2003a). Operator functional state classification using multiple psychophysiological features in an air traffic control task. [Article]. *Human Factors, 45*(3), 381-389.

Wilson, G. F., & Russell, C. A. (2003b). Real-time assessment of mental workload using psychophysiological measures and artificial neural networks. *Human Factors, 45*(4), 635-643.

Wilson, G. F., & Russell, C. A. (2007). Performanee enhancement in an uninhabited air vehiele task using psychophysiologically determined adaptive aiding. *Human Factors, 49*(6), 1005-1018.

Yu, Zhaohan. (2009). *OPTIMIZATION TECHNIQUES IN DATA MINING WITH APPLICATIONS TO BIOMEDICAL AND PSYCHOPHYSIOLOGICAL DATA SETS.* University of Iowa, Iowa City.

Zhang, D. D., Ding, H. Y., Liu, Y. F., Zhou, C., Ding, H. S., & Ye, D. T. (2009). Neurodevelopment in newborns: a sample entropy analysis of electroencephalogram. *Physiological Measurement, 30*(5), 491-504.